

=> fil hcaplus
FILE 'HCAPLUS' ENTERED AT 15:31:14 ON 04 MAR 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

FILE COVERS 1907 - 4 Mar 2003 VOL 138 ISS 10
FILE LAST UPDATED: 3 Mar 2003 (20030303/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

=> d que 120

L1 8 SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTGDLVTYDKENGMHKKVFYSFIDD
KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDYYPNRSID
TK.Y.S.L.Y.F.G.V.G.D.G.I.G.I.A.V.I.H.L.Y.QPDFKFILESPTDKKVGWKVIF
NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP
L2 6 SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTGDLVTYDKENGMHKKVFYSFIDD
KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDYYPNRSID
T.E.M.T.T.G.N.N.T.D.T.K.G.L.G.N.S.G.T.K.VQPDFKFILESPTDKKVGWKVIF
NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP
L3 8 SEA FILE=REGISTRY L1 OR L2
L4 3 SEA FILE=HCAPLUS L3
L5 1 SEA FILE=HCAPLUS L4 AND PORE#
L6 1 SEA FILE=HCAPLUS L4 AND STEM#
L7 1 SEA FILE=HCAPLUS L4 AND LUMEN?
L8 1 SEA FILE=HCAPLUS L4 AND TRANSMEMBRANE(3A)CHANNEL?
L9 0 SEA FILE=HCAPLUS L4 AND HEPTAMER?
L10 1 SEA FILE=HCAPLUS L4 AND HETEROLOG?
L11 1 SEA FILE=HCAPLUS L4 AND (SENSOR? OR BIOSENSOR?)
L12 1 SEA FILE=HCAPLUS L4 AND METAL
L13 0 SEA FILE=HCAPLUS L4 AND CURRENT?
L14 0 SEA FILE=HCAPLUS L4 AND DIGITAL
L20 1 SEA FILE=HCAPLUS (L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11 OR
L12 OR L13 OR L14)

=> d ibib abs 120 1

L20 ANSWER 1 OF 1 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1999:96264 HCAPLUS
DOCUMENT NUMBER: 130:165159
TITLE: Designed staphylococcal hemolysin protein
pores as components for metal
biosensors
INVENTOR(S): Bayley, Hagan; Braha, Orit; Kasianowicz, John; Gouaux,
Eric

THIS PAGE BLANK (USPTO)

PATENT ASSIGNEE(S): University of Massachusetts, USA
 SOURCE: PCT Int. Appl., 51 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9905167	A1	19990204	WO 1998-US15354	19980724
W: AU, CA, JP, KR				
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
AU 9885862	A1	19990216	AU 1998-85862	19980724
PRIORITY APPLN. INFO.:			US 1997-53737P	P 19970725
			WO 1998-US15354	W 19980724

AB This invention features a mutant staphylococcal alpha hemolysin (.alpha.HL) polypeptide contg. a **heterologous metal**-binding amino acid. The polypeptide assembles into a heteroheptameric **pore** assembly in the presence of a wild type .alpha.HL polypeptide. Preferably, the **metal**-binding amino acid occupies a position in a **transmembrane channel** of the heteroheptameric **pore** assembly, e.g., an amino acid in the **stem** domain of WT .alpha.HL is substituted with a **heterologous metal**-binding amino acid. More preferably, the **metal**-binding amino acid projects into the **lumen** of the **transmembrane channel**.

REFERENCE COUNT: 1 THERE ARE 1 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

=> d que 122

L1 8 SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTGDLVTYDKENGMHKKVFYFSFIDD
 KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDIYPRNSID
 TK.Y.S.L.Y.F.G.V.G.D.G.I.G.I.A.V.I.H.L.Y.QPDFKFILESPTDKKVGWKVIF
 NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP

L2 6 SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTGDLVTYDKENGMHKKVFYFSFIDD
 KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDIYPRNSID
 T.E.M.T.T.G.N.N.T.D.T.K.G.L.G.N.S.G.T.K.VQPDFKFILESPTDKKVGWKVIF
 NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP

L3 8 SEA FILE=REGISTRY L1 OR L2

L4 3 SEA FILE=HCAPLUS L3

L5 1 SEA FILE=HCAPLUS L4 AND PORE#

L6 1 SEA FILE=HCAPLUS L4 AND STEM#

L7 1 SEA FILE=HCAPLUS L4 AND LUMEN?

L8 1 SEA FILE=HCAPLUS L4 AND TRANSMEMBRANE (3A) CHANNEL?

L9 0 SEA FILE=HCAPLUS L4 AND HEPTAMER?

L10 1 SEA FILE=HCAPLUS L4 AND HETEROLOG?

L11 1 SEA FILE=HCAPLUS L4 AND (SENSOR? OR BIOSENSOR?)

L12 1 SEA FILE=HCAPLUS L4 AND METAL

L13 0 SEA FILE=HCAPLUS L4 AND CURRENT?

L14 0 SEA FILE=HCAPLUS L4 AND DIGITAL

L15 640 SEA FILE=HCAPLUS STAPH? (3A) (HEMOLYSIN# OR HAEMOLYSIN#)

L16 6 SEA FILE=HCAPLUS L15 AND ((SUBSTITUT? OR HETEROLOG? OR
 REPLAC?) (3A) (AMINO(A)ACID#))

L17 3 SEA FILE=HCAPLUS (L16 AND (PORE# OR STEM# OR LUMEN? OR
 TRANSMEMBRANE (3A) CHANNEL?))

L18 2 SEA FILE=HCAPLUS L16 AND (SENSOR? OR BIOSENSOR?)

THIS PAGE BLANK (USPTO)

L19 2 SEA FILE=HCAPLUS L16 AND (METAL# OR CURRENT# OR DIGITAL)
L20 1 SEA FILE=HCAPLUS (L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11 OR
L12 OR L13 OR L14)
L21 3 SEA FILE=HCAPLUS (L17 OR L18 OR L19)
L22 2 SEA FILE=HCAPLUS L21 NOT L20

=> d ibib abs l22 1-2

L22 ANSWER 1 OF 2 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1997:331282 HCAPLUS

DOCUMENT NUMBER: 127:62190

TITLE: A predicted .beta.-sheet from class S components of
staphylococcal .gamma.-hemolysin is
essential for the secondary interaction of the class F
component

AUTHOR(S): Meunier, O.; Ferreras, M.; Supersac, G.; Hoeper, F.;
Baba-Moussa, L.; Monteil, H.; Colin, D. A.;
Menestrina, G.; Prevost, G.

CORPORATE SOURCE: Institut de Bacteriologie de la Faculte de Medecine de
Strasbourg, 3, rue Koeberle, Strasbourg, F-67000, Fr.

SOURCE: Biochimica et Biophysica Acta (1997), 1326(2), 275-286
CODEN: BBACAQ; ISSN: 0006-3002

PUBLISHER: Elsevier

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Site-directed mutagenesis was performed on genes encoding HlgA and HlgC,
two of the three proteins expressed from the **staphylococcal**
.gamma.-**hemolysin** locus, which originate two **pore**
-forming toxins (HlgA+HlgB, HlgC+HlgB). As related proteins, HlgA and
HlgC were found to bind first to cell membranes. **Amino**
acid substitutions concerning residues that would
predictably disrupt a 13 amino acid conserved .beta.-sheet of the Chou and
Fasman secondary structure prediction were created. The mutation of a
threonine into an aspartic acid residue from HlgA (T28D) and from HlgC
(T30D) that would break this predicted N-terminal structure lowered
dramatically the biol. activities towards purely lipidic vesicles,
erythrocytes, and polymorphonuclear cells. The change in secondary
structure was confirmed by Fourier Transform IR spectroscopy. The binding
of mutated and native proteins at the same kind of sites onto
polymorphonuclear cells was evidenced with flow cytometry and
fluorescein-labeled anti-class S antibodies or wild type HlgA or HlgC.
However, the subsequent binding of fluorescein-labeled HlgB to
membrane-bound mutated HlgA or HlgC complexes was inhibited. In
conclusion, the initial binding of class S components is essential for the
subsequent binding of class F components, and a predicted .beta.-sheet
seems to be at least one of the functional domains involved.

L22 ANSWER 2 OF 2 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1994:529301 HCAPLUS

DOCUMENT NUMBER: 121:129301

TITLE: A **pore**-forming protein with a **metal**
-actuated switch

AUTHOR(S): Walker, Barbara; Kasianowicz, John; Krishnasastri,
Musti; Bayley, Hagan

CORPORATE SOURCE: Worcester Found. Exp. Biol., Shrewsbury, MA, 01545,
USA

SOURCE: Protein Engineering (1994), 7(5), 655-62
CODEN: PRENE9; ISSN: 0269-2139

THIS PAGE BLANK (USPTO)

DOCUMENT TYPE: Journal
 LANGUAGE: English

AB **Staphylococcal .alpha.-hemolysin, a pore**
 -forming exotoxin, is a polypeptide of 293 amino acids that is secreted by *Staphylococcus aureus* as a water-sol. monomer. It assembles to form hexameric **pores** in lipid bilayers. Previous studies of **pore** formation have established the involvement of a central glycine-rich loop. Here, the authors show that when five consecutive histidine residues **replace amino acids** 130-134 at the midpoint of the loop, they provide a switch with which **pore** activity can be (i) turned off by micromolar concns. of divalent zinc ions and (ii) turned back on with the chelating agent EDTA. Planar bilayer recordings show that Zn²⁺ and EDTA can act on open channels from either side of the bilayer and thus demonstrate that the central loop lines part of the conductive pathway. The authors' results suggest that genetically-engineered **pore**-forming proteins might make useful components of **metal ion sensors**.

=> d que 123

L1 8 SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTDGLVITYDKENGMHKKVFYSFIDD
 KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDYYPNSID
 TK.Y.S.L.Y.F.G.V.G.D.G.I.G.I.A.V.I.H.L.Y.QPDFKFILESPTDKKVGWKVIF
 NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP
 L2 6 SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTDGLVITYDKENGMHKKVFYSFIDD
 KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDYYPNSID
 T.E.M.T.T.G.N.N.T.D.T.K.G.L.G.N.S.G.T.K.VQPDFKFILESPTDKKVGWKVIF
 NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP
 L3 8 SEA FILE=REGISTRY L1 OR L2
 L4 3 SEA FILE=HCAPLUS L3
 L5 1 SEA FILE=HCAPLUS L4 AND PORE#
 L6 1 SEA FILE=HCAPLUS L4 AND STEM#
 L7 1 SEA FILE=HCAPLUS L4 AND LUMEN?
 L8 1 SEA FILE=HCAPLUS L4 AND TRANSMEMBRANE (3A) CHANNEL?
 L9 0 SEA FILE=HCAPLUS L4 AND HEPTAMER?
 L10 1 SEA FILE=HCAPLUS L4 AND HETEROLOG?
 L11 1 SEA FILE=HCAPLUS L4 AND (SENSOR? OR BIOSENSOR?)
 L12 1 SEA FILE=HCAPLUS L4 AND METAL
 L13 0 SEA FILE=HCAPLUS L4 AND CURRENT?
 L14 0 SEA FILE=HCAPLUS L4 AND DIGITAL
 L20 1 SEA FILE=HCAPLUS (L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11 OR
 L12 OR L13 OR L14)
 L23 2 SEA FILE=HCAPLUS L4 NOT L20

=> d ibib abs 123 1-2

L23 ANSWER 1 OF 2 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2002:906293 HCAPLUS

DOCUMENT NUMBER: 138:8311

TITLE: *Staphylococcus aureus* proteins and nucleic acids and
 their diagnostic and therapeutic uses for
 staphylococcal infections

INVENTOR(S): Masignani, Vega; Mora, Marirosa; Scarselli, Maria

PATENT ASSIGNEE(S): Chiron Spa, Italy

SOURCE: PCT Int. Appl., 49 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

THIS PAGE BLANK (USPTO)

FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002094868	A2	20021128	WO 2002-IB2637	20020327
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				
PRIORITY APPLN. INFO.:			GB 2001-7661	A 20010327
AB The invention provides 2821 nucleic acid coding sequences from Staphylococcus aureus strain NCTC 8325 along with their inferred translation products. The proteins are useful for vaccines, immunogenic compns., diagnostics, enzymic studies, and also as targets for antibiotics.				

L23 ANSWER 2 OF 2 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1995:883755 HCAPLUS

DOCUMENT NUMBER: 124:23160

TITLE: Activation of .alpha.-toxin translation in
Staphylococcus aureus by the trans-encoded antisense
RNA, RNAIII

AUTHOR(S): Morfedt, Eva; Taylor, David; von Gabain, Alexander;
Arvidson, Staffan

CORPORATE SOURCE: Microbiol. Tumoriol. Cent., Karolinska Inst.,
Stockholm, S-171 77, Swed.

SOURCE: EMBO Journal (1995), 14(18), 4569-77
CODEN: EMJODG; ISSN: 0261-4189

PUBLISHER: Oxford University Press

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The synthesis of virulence factors in Staphylococcus aureus is controlled by a regulatory RNA mol., RNAIII, encoded by the agr locus. Transcription of genes coding for secreted toxins and enzymes is stimulated, while transcription of cell-surface protein genes is repressed by RNAIII. In the case of staphylococcal .alpha.-toxin, RNAIII also seems to stimulate translation by an independent mechanism. In this report we show that in a mutant lacking RNAIII the rate of .alpha.-toxin (hla) prodn. relative to the cellular concn. of hla mRNA was reduced 10-fold as compared with the wild-type strain. A 75% complementarity between the 5' end of RNAIII and the 5' untranslated region of the hla transcript suggests a direct interaction between the RNAs. A complex of RNAIII and hla mRNA was demonstrated in exts. of total RNA from the wild-type strain, and also with in vitro synthesized RNAs. RNase T1 digestion expts. revealed that the ribosome binding site of the hla transcript is blocked by intramol. base-pairing. Hybridization with RNAIII prevents this intramol. base-pairing and makes the hla mRNA accessible for translation initiation. This is, to our knowledge, the first example of an antisense RNA that stimulates translation of the target mRNA.

THIS PAGE BLANK (08PTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: March 4, 2003, 10:38:18 ; Search time 33 Seconds
(without alignments)

1183.104 Million cell updates/sec

Title: CLAIM6

Perfect score: 1485

Sequence: 1 ADSINIKTGTDTGSNTTV.....WTDSSERYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1485	100.0	293	20	AAV01365
2	424.5	28.6	330	22	AAU33797
3	424.5	28.6	333	22	AAU33797
4	404	27.2	325	22	AAU334287
5	404	27.2	325	22	AAU334287
6	404	27.2	325	22	AAU37100
7	404	27.2	325	22	AAU37186
8	404	27.2	325	22	AAU37477
9	403.5	27.2	327	22	AAU34405
10	403.5	27.2	327	22	AAU37136

11	403.5	27.2	327	22	AAU37564	Staphylococcus aur
12	397	26.7	323	22	AAU34395	Staphylococcus aur
13	397	26.7	323	22	AAU34395	Staphylococcus aur
14	390	26.3	323	14	AAU37509	(Pro)leukocidin F.
15	252	17.0	321	22	AAU34275	Staphylococcus aur
16	252	17.0	321	22	AAU37101	Staphylococcus aur
17	252	17.0	321	22	AAU37128	Staphylococcus aur
18	238	16.0	315	14	AAU35083	(Pro)leukocidin S.
19	234	15.8	312	22	AAU33771	Staphylococcus aur
20	126	8.5	40	20	AAU01366	Alpha haemolysin (
21	124.5	8.4	826	22	AAU01366	S. epidermidis ope
22	124.5	8.4	10182	23	ABP38314	Staphylococcus epi
23	115.5	7.8	1233	21	AAU18246	Plasmidium falcipa
24	115	7.7	157	23	AAU75479	S. aureus antigeni
25	114.5	7.7	674	22	AAU92775	Human protein sequ
26	113.5	7.6	550	23	ABU53650	Lactococcus lactis
27	113	7.6	435	23	ABU55618	Lactococcus lactis
28	111	7.5	461	14	AAU39354	EpiP protein. Sta
29	111	7.5	461	20	AAU43441	S. epidermis readi
30	110.5	7.4	1231	20	AAU00219	Enterococcus faeca
31	110.5	7.4	1231	23	ABP43438	E faecalis EF108 a
32	110.5	7.4	1265	20	AAU00218	Enterococcus faeca
33	110.5	7.4	1265	23	ABP43437	Enterococcus faeca
34	109.5	7.4	941	22	ABG25801	Novel human diagno
35	109.5	7.4	1300	22	AAU33407	Enterococcus faeca
36	109.5	7.4	1300	22	AAU35313	Enterococcus faeca
37	108.5	7.3	287	18	AAU18559	Novel fusion prote
38	108.5	7.3	918	20	AAU08640	S. aureus ClfB pro
39	108.5	7.3	1487	16	AAU74634	Bacillus circulans
40	108	7.3	387	10	AAU90119	Heat stable sarcos
41	108	7.3	1023	22	ABG22883	Novel human diagno
42	108	7.3	1023	22	AAU79772	Human protein SEQ
43	108	7.3	1023	22	AAU79773	Human protein SEQ
44	108	7.3	1076	23	AAU74355	Human cytoskeleton
45	108	7.3	1227	22	AAU65637	Novel protein kina

ALIGNMENTS

RESULT 1

AAU01365
ID AAY01365 standard; protein; 293 AA.

XX AAY01365;

XX 03-JUN-1999 (first entry)

XX Wild-type Staphylococcal alpha haemolysin (HL) polypeptide.

XX Staphylococcal; alpha-haemolysin; alphaHL; mutant; metal; biosensor;
XX heptameric pore assembly; micronutrient analysis; industrial effluent;
XX organic compound; explosive; macromolecule; bacteria; virus.

XX Staphylococcus aureus.

XX WO905167-A1.

XX 04-FEB-1999.

XX 24-JUL-1998; 98WO-US15354.

XX 25-JUL-1997; 97US-0053737.

XX (UYMA-) UNIV MASSACHUSETTS.

XX Bayley H, Braha O, Gouaux E, Kasianowicz J;

XX WPI; 1999-153311/13.

XX New mutant staphylococcal alpha-haemolysin - comprises a
XX heterologous amino acid that binds to analyte, particularly metal
XX ions

XX Disclosure; Page 12; 51pp; English.

XX The invention relates to a mutant staphylococcal alpha-haemolysin

CC (alphaHL) polypeptide. The mutant alphaHL includes a heterologous amino

CC acid (HAA) that binds an analyte, and wherein the polypeptide assembles

CC into a heteroheptameric pore assembly in presence of many wild-type

CC alphaHL polypeptides. The mutant alphaHL polypeptide has at least two

CC non-consecutive HAA in its stem domain, each of which binds: (i) a metal,

CC or (ii) an organic compound. The HAA occupy two or more of the amino acid

CC positions selected from residues 110-149 of the wild-type alphaHL (the

CC present sequence). The HAA can be selected from the group consisting of

CC (i) Ser, Thr, Met, Trp and Tyr or (ii) Glu, Asp, Cys or especially His.

CC Digital biosensors (comprising heptameric pore assemblies containing the

CC mutant alphaHL) are particularly used to detect and quantify metal ions

CC (specifically zinc, cobalt, nickel and cadmium), e.g. in water (for

CC micronutrient analysis), sediment, air, industrial effluent. Organic

CC compounds that can be detected are specifically explosives, but may also

CC be macromolecules or entire bacteria or viruses. The present sequence

CC represents a wild-type staphylococcal alpha haemolysin polypeptide.

XX Sequence 293 AA;

XX

Query Match 100.0%; Score 1485; DB 20; Length 293;

Best Local Similarity 93.2%; Pred. No. 1.5e-113;

Matches 273; Conservative 20; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADSDINIKTGTDTIGSNTTVKTDGVTYDKENGMMHKVKVYFSDIDKNNHKKLLVIRTKGT 60

DB 1 ADSDINIKTGTDTIGSNTTVKTDGVTYDKENGMMHKVKVYFSDIDKNNHKKLLVIRTKGT 60

QY 61 IAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSXLXYXF 120

DB 61 IAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSXLXYXF 120

QY 121 XGVYXGDXGXIXGXIXAXVXIXHXLXYXQDPFKTILESPT-DKKVGKVKVFNMMVQNW 180

DB 121 NGNVTGDDTKIGKGLIGANVSIGHTLUKYVQDPFKTILESPTDKKVGKVKVFNMMVQNW 180

QY 181 PYDRDSWNPVYGQNFPMKTRNGSMKAADNFDLPNKAASSLLSGFSPDFATVITMDRKASK 240

DB 181 PYDRDSWNPVYGQNFPMKTRNGSMKAADNFDLPNKAASSLLSGFSPDFATVITMDRKASK 240

QY 241 QQTNDIVYVRVDDYQLHWTSTNNKGTNTKDKWTDSSERYKIDWEKEEMTN 293

DB 241 QQTNDIVYVRVDDYQLHWTSTNNKGTNTKDKWTDSSERYKIDWEKEEMTN 293

RESULT 2

ID AAU33797

XX AAU33797 standard; Protein; 330 AA.

XX

XX AAU33797;

XX

XX 14-FEB-2002 (first entry)

XX

XX Staphylococcus aureus cellular proliferation protein #73.

XX

XX Antisense; prokaryotic cellular proliferation protein;

XX antibiotic; antibacterial; drug design.

XX

XX Staphylococcus aureus.

XX

XX WO200170955-A2.

XX

XX 27-SEP-2001.

XX

XX 21-MAR-2001; 2001WO-US09180.

XX

XX 21-MAR-2000; 2000US-191078P.

XX

XX 23-MAY-2000; 2000US-206848P.

XX

XX 26-MAY-2000; 2000US-207727P.

XX

XX 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX

XX WPI: 2001-611495/70.

DR N-PSDB; AA551656.

XX

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids -

XX

XX Example 3; Seq ID No 5293; 51pp; English.

XX

XX The invention relates to antisense inhibitors of genes essential to

XX prokaryotic cellular proliferation, their use in identifying the

XX genes, their use in the discovery of novel antibiotics, the essential

XX genes themselves and the encoded proteins. The prokaryotes used are

XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

XX invention is also useful for the identification of potential new targets

XX for antibiotic development. The antisense nucleic acids can also be used

XX to identify proteins used in proliferation, to express these proteins,

XX and to obtain antibodies capable of binding to the expressed proteins.

XX The proteins can be used to screen compounds in rational drug discovery

XX programmes. The antisense nucleic acid sequence is also useful to screen

XX for homologous nucleic acids which are required for cell proliferation in

XX a wide variety of organisms. The present sequence represents an

XX essential prokaryotic cellular proliferation protein.

XX Note: The sequence data for this patent did not form part

XX of the printed specification, but was obtained in electronic

XX format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 330 AA;

Query Match 28.6%; Score 424.5; DB 22; Length 330;

Best Local Similarity 30.8%; Pred. No. 5e-26;

Matches 91; Conservative 80; Mismatches 107; Indels 17; Gaps 8;

QY 7 IKGTGTD-----IGSNTTVKTDGVTYDKENGMMHKVKVYFSDIDKNNHKKLLVIRTKGTI 61

DB 30 IKTEITPVVEKAIDSETKMYTATSTDEKNISQSLQFNFLNEPNYDKETLFIKAKGTI 89

QY 62 AGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSXLXYXF 120

DB 90 SSGLEILYPNGYWSSTLRWPGSYSVSIQNVDDNTTKVTDFAPKNQDTRDVKYTYGKT 149

QY 121 XGVYXGDXGXIXGXIXAXVXIXHXLXYXQDPFKTILESPT-DKKVGKVKVFNMMVQNW 179

DB 150 GGNFT-INQNGLSGNITQYNYSETISYQQPSYRTLIDKPTSNKAVAMKVEANVKNGY 208

QY 180 GPYDRDS---WNPVYGQNFPMKTRNGSMKAADNFDLPNKAASSLLSGFSPDFATVITMDR 236

DB 209 -DHTRDATDGTGKVGSEIFLTRNGNLWARDNFTPKKMPVTVSEGNPEFLTVMHSHDK 267

QY 237 KASKOQTNDIVYVRVDDYQLHWTSTN---WKGNTKDKWTDSSERYKIDWE 287

DB 268 K-KGKHSLFVVYRKVRMDDFEIRWNYRNWGYWGSCKNNVKNKKEESLAALYEIDWK 321

RESULT 3

ID AAU36528

XX AAU36528 standard; Protein; 333 AA.

XX

XX AAU36528;

XX

XX 14-FEB-2002 (first entry)

XX

XX Staphylococcus aureus cellular proliferation protein #698.

DR WPI: 2001-611495/70.
XX N-PSDB; AAS55045.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12779; 51lpp; English.
XX The invention relates to antisense inhibitors of genes essential
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes, themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 325 AA;

Query Match 27.2%; Score 404; DB 22; Length 325;
Best Local Similarity 29.4%; Pred. No. 2.4e-24;
Matches 87; Conservative 74; Mismatches 119; Indels 16; Gaps 6;

QY 9 TGTTDI-GSNTTV---KTGDLVYDKENG-----MHKKVYFSDDKNNHKKLLVIR 56
Db 21 SGTANAEGKITPVSVKVKDDKVTLYKTATADSKFKISQILTFNFKDKSYDKDTLVLK 80

QY 57 TKGTIAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTKKXXSXL 116
Db 81 ATGNINSFGVKPNPDYDFSKLWAGAKYNYVSISSQSDNSVNVVDYAPKNQNEEFQVQNTL 140

QY 117 XYFXGXVXGDXGXIXGXIXAXVXIXHXLXYQDPDKTILESPTD-KKVGWKVIFNNMV 175
Db 141 GYTFGGDI--SISNGLSGGLNGTAFSETINYKQESYRTTLRSNTYKNVGVGVEAHKIM 198

QY 176 NQWNGPYDRDSWNPVYGNQLFMKTRNGSMKAADNFDLPNKASSLLSGSFPDFATVITMD 235
Db 199 NNGWGPYGRDSFHTYGNELFLAGRQSSAYAGONFIAQHQMPLLSRSNFPNPEFLSVLS-H 257

QY 236 RKASKQOTNIDVIYVRVDDYQLHWTSTNNKGTNTKDKWDRSSERYKIDWEKEEM 291
Db 258 RDGAKKSKITVYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 313

RESULT 7
AAU37477
ID AAU37477 standard; Protein; 325 AA.
XX
XX AAU37477;
XX
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1647.
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
OS
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
DR N-PSDB; AAS55336.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 13070; 51lpp; English.
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes, themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 325 AA;

Query Match 27.2%; Score 404; DB 22; Length 325;
Best Local Similarity 29.4%; Pred. No. 2.4e-24;
Matches 87; Conservative 74; Mismatches 119; Indels 16; Gaps 6;

QY 9 TGTTDI-GSNTTV---KTGDLVYDKENG-----MHKKVYFSDDKNNHKKLLVIR 56
Db 21 SGTANAEGKITPVSVKVKDDKVTLYKTATADSKFKISQILTFNFKDKSYDKDTLVLK 80

QY 57 TKGTIAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTKKXXSXL 116
Db 81 ATGNINSFGVKPNPDYDFSKLWAGAKYNYVSISSQSDNSVNVVDYAPKNQNEEFQVQNTL 140

QY 117 XYFXGXVXGDXGXIXGXIXAXVXIXHXLXYQDPDKTILESPTD-KKVGWKVIFNNMV 175
Db 141 GYTFGGDI--SISNGLSGGLNGTAFSETINYKQESYRTTLRSNTYKNVGVGVEAHKIM 198

QY 176 NQWNGPYDRDSWNPVYGNQLFMKTRNGSMKAADNFDLPNKASSLLSGSFPDFATVITMD 235
Db 199 NNGWGPYGRDSFHTYGNELFLAGRQSSAYAGONFIAQHQMPLLSRSNFPNPEFLSVLS-H 257

QY 236 RKASKQOTNIDVIYVRVDDYQLHWTSTNNKGTNTKDKWDRSSERYKIDWEKEEM 291
Db 258 RDGAKKSKITVYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 313

RESULT 8
AAU75478


```
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
DR N-PSDB; AAS55423.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT
XX Example 3; Seq ID No 13157; 51lpp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 327 AA:
XX
XX Query Match 27.2%; Score 403.5; DB 22; Length 327;
XX Best Local Similarity 28.8%; Pred. No. 2.7e-24;
XX Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;
XX
QY 18 TTVKTDGLVTDKMGHKKVYFSDDKNNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
DB 43 TLYKTTATSDNKLN-ISQILFFNIKDKSYDKDITLVKAAAGNINSYKPKPNKYNSQ 101
QY 78 LAWPSAFKVLQLPDNEVAQISDYYPNRSIDTKYXSXLXYFXGXGXDXGXIXGXIX 137
DB 102 FYWGGKYNVSSESNDVAVVYAPKQNEFFVQQTILGYSYGGDI--NISNGLSGGLN 159
QY 138 AXVXIXHXLXYXQPFKTILESPTD-KVGVKVIFFNNMVNQNMGYPYDRSDMNPVYGNOLF 196
DB 160 GSKSFSETINYKQESYRTTIDRKTNHSKIGWGVEAHKIMNNGWGPYGRSDYDPTVGNELF 219
QY 197 MKTRNGSKAADNFDLPNKASLLSSGSPDPATVITWDRKASKOQTNDIVYIERVDDY 256
DB 220 LGGRSSNAGCNFLPTHTOMPLLAGRNENPEFISVLSHKQNDTK-KSKIKVTYQREMDRY 278
QY 257 QLHWTSTNKGKNTKDKWTDSSERYKIDWE 287
DB 279 TNQWRNLHWGNNYKNQNTVFTSYEVDWQ 309
XX
XX RESULT 12
XX AAU34395
XX ID AAU34395 standard; Protein; 323 AA.
XX
XX AAU34395;
XX AC
XX
XX 14-FEB-2002 (first entry)
```

```
XX
DE Staphylococcus aureus cellular proliferation protein #671.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
XX
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
XX 23-MAY-2000; 2000US-208848P.
XX 26-MAY-2000; 2000US-207727P.
XX 23-OCT-2000; 2000US-242578P.
XX 27-NOV-2000; 2000US-253625P.
XX 22-DEC-2000; 2000US-257931P.
XX 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
XX N-PSDB; AAS52254.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 5891; 51lpp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the
XX genes, their use in the discovery of novel antibiotics, the essential
XX genes themselves and the encoded proteins. The prokaryotes used are
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX invention is also useful for the identification of potential new targets
XX for antibiotic development. The antisense nucleic acids can also be used
XX to identify proteins used in proliferation, to express these proteins,
XX and to obtain antibodies capable of binding to the expressed proteins.
XX The proteins can be used to screen compounds in rational drug discovery
XX programmes. The antisense nucleic acid sequence is also useful to screen
XX for homologous nucleic acids which are required for cell proliferation in
XX a wide variety of organisms. The present sequence represents an
XX essential prokaryotic cellular proliferation protein.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 323 AA;
XX
XX Query Match 26.7%; Score 397; DB 22; Length 323;
XX Best Local Similarity 28.5%; Pred. No. 8.9e-24;
XX Matches 81; Conservative 83; Mismatches 108; Indels 12; Gaps 7;
XX
QY 13 DIGSNTTVKTDGLVTDKMGHKKVYFSDDKNNHKKLLVIRTKGTIAGQYRVYSEEG 72
DB 36 NLDGDKMYTTRATSDSQKNITQSLQNFLETPEYDKETVFIKAKGHIGSLRILDPNG 95
QY 73 ANKSLGAWPSAFKVLQ-LPDNEVAQISDYYPNRSIDTKYXSXLXYFXGXGXDXGX 131
DB 96 YWNSTLRWPGSYSVSIQVNDNNNTNTVDFAPKNODESREVKYTVGYKTGDD-PSINRG 154
QY 132 IXGXIXAXVXIXHXLXYXQPFKTILESPTDK-KVGVKVIFFNNMVNQNMGYPYDRSD 187
DB 155 LTGNITKESYSETISYQOQPSYRTLLDQSTSHKGVGWKVEAHKIMNNGHDTHTQLTNDSD 214
```

QY 188 NPYVGNQFMKTRNGSMKAADNFDLPNKASSLLSSGFSDFATVITMDRKASQOQNIDV 247
 Db 215 NRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTYSEGFNPEFLAVMSHDKK-DRGKSOQFVV 272
 QY 248 IYERVRDDYQLHWTSTN---WKGNTKDKWTDSSERYKIDWE 287
 Db 273 HYKRSWDEFKIDWNHRHGFYSGENHVDKKEKLSALYEVDWK 316
 RESULT 13
 AAU37509
 ID AAU37509 standard; Protein; 338 AA.
 XX
 AC AAU37509;
 DT 14-FEB-2002 (first entry)
 XX
 Staphylococcus aureus cellular proliferation protein #1679.
 DE Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 KW Staphylococcus aureus.
 OS
 XX WO200170955-A2.
 PN 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS55368.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 13102; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 338 AA;

Query Match 26.7%; Score 397; DB 22; Length 338;
 Best Local Similarity 28.5%; Pred. No. 9.8e-24;
 Matches 81; Conservative 83; Mismatches 108; Indels 12; Gaps 7;
 QY 13 DIGSTNTVKTGDLVTDKENGHKKVPYSPFIDDKNNKKLLVIRTKGTIAGQYRVISEEG 72
 Db 42 NLDGDKMVTTRTATSDSQKNITQSLQFNPLTEPNYDKETVFKAKGTIGSLRILDPNG 101
 QY 73 ANKSGLAWPSAFKVLQ-LPDNEVAOISDYPRNSIDTKYXSXLXYYFXGVXGXDXGX 131
 Db 102 YWNSTLRWPGSYSVSIQNDNNNTNFTDPAKNQDSREVKYTYGKTGGD-FSINRGG 160
 QY 132 IXGXIXAXVXIXHXLXYXOPDFKTIILESPDKK-VGKVV---IFNNMVNNGWGPYDRDSW 187
 Db 161 LTGNITKESNYSETISYQPSYRTLLDOSTSHKGVGKVEAHLLNNMGHDHTQLTNDSD 220
 QY 188 NPYVGNQFMKTRNGSMKAADNFDLPNKASSLLSSGFSDFATVITMDRKASQOQNIDV 247
 Db 221 NRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTYSEGFNPEFLAVMSHDKK-DRGKSOQFVV 278
 QY 248 IYERVRDDYQLHWTSTN---WKGNTKDKWTDSSERYKIDWE 287
 Db 279 HYKRSWDEFKIDWNHRHGFYSGENHVDKKEKLSALYEVDWK 322
 RESULT 14
 AAR35084
 ID AAR35084 standard; Protein; 323 AA.
 XX
 AC AAR35084;
 DT 09-AUG-1993 (first entry)
 DE (Pro)leukocidin F.
 XX
 KW Detection; methicillin-resistant Staphylococcus aureus;
 KW leukocidin; proleukocidin; cancer; antibody.
 XX
 OS Methicillin-resistant Staphylococcus strain 4 (RIMD-3109025).
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..25 /label= sig_peptide
 FT Protein 26..323 /label= mat_protein
 FT
 XX WO9307487-A.
 XX
 PD 15-APR-1993.
 XX
 PF 09-OCT-1992; 92WO-JP01317.
 XX
 PR 09-OCT-1991; 91JP-0290742.
 PR 23-APR-1992; 92JP-0129956.
 XX
 PA (IATR) IATRON LAB INC.
 XX
 PI Izaki K, Kamio Y, Kita H;
 XX
 DR WPI; 1993-134623/16.
 DR N-PSDB; AAQ40216.
 XX
 PT Early detection of methicillin-resistant Staphylococcus aureus -
 PT using antibodies to leukocidin and proleukocidin, useful for
 PT cancer patients, elderly etc.
 XX
 PS Disclosure; Fig 13; 67pp; Japanese.
 XX
 CC DNA and corresp. amino acid sequences of (pro)leukocidin S and F
 CC are claimed. A novel method for the detection of
 CC methicillin-resistant Staphylococcus aureus (MRSA) in biological
 CC samples comprises reaction with an antibody which recognises
 CC (pro)leukocidin S or F.

XX	SQ	Sequence	323 AA;
		Query Match	26.3%; Score 390; DB 14; Length 323;
		Best Local Similarity	29.2%; Pred. No. 3.4e-23;
		Matches	86; Conservative 73; Mismatches 120; Indels 16; Gaps 6;
Qy	9	TGTTDI-GSNTV---KTGDVLVTYDKENG-----MHKKVFYSFIDDKNHNKLLLVIR	56
Dd	21	SGTAEAGKITPVSVKKVDVKTYLTATADSKFKISQLTFENFIKDKEYDXTLVLK	80
Qy	57	TKGTIACQRYVSEEGANKSLWAPSAFKVOLPLDPNEVAQISDYPRNSITDTKYXSKL	116
Dd	81	ATGNINSFGVPKNPNDFESKLWGAKYNVISISSQSNDVNADVAPKQNKEEFQVQNTL	140
Qy	117	XYFXGVGXVGDXDGXIAXGXIXAIVKHXLXYPXDPTILESPTDKKVGWKVIFNNMVN	176
Dd	141	GYTEGGDI--SISNLGGGLNGNTAFSETINYOESTRTLSRNNYKNVGMGVEAHKIMN	198
Qy	177	QNWGPYDRDSNPVYGQLFMKTRTGSKMGKAADFDPNKASSLLSSGSFPDAFTVITMDR	236
Dd	199	-GWGPFQRDSFHPTYGNELFLAQRSSAYAGQNFIAHQHMPLLSRSNFNPRELSVLSHRQ	257
Qy	237	KASQQQTNIIDVIYERVRDDYOLHWTSNWKGTNTKWTDRESSBRYKIDWEKEM	291
Dd	258	DAAR-KSKITVTYQREMDLYQIRWNGFYWAGANYKFNRTRTFKSTYEIDWNHKV	311
 RESULT 15 AAU34275			
ID	AAU34275	standard; Protein; 321 AA.	
XX	AAU34275;		
XX			
DT	14-FEB-2002	(first entry)	
XX		Staphylococcus aureus cellular proliferation protein #551.	
DE		Antisense; prokaryotic cellular proliferation protein;	
KW		antibiotic; antibacterial; drug design.	
XX			
OS		Staphylococcus aureus.	
XX			
PN	WO200170955-A2.		
XX			
PD	27-SEP-2001.		
XX			
PF	21-MAR-2001; 2001WO-US09180.		
XX			
PR	21-MAR-2000; 2000US-191078P.		
PR	23-MAY-2000; 2000US-206848P.		
PR	26-MAY-2000; 2000US-207727P.		
PR	23-OCT-2000; 2000US-242578P.		
PR	27-NOV-2000; 2000US-253625P.		
PR	22-DEC-2000; 2000US-257931P.		
PR	16-FEB-2001; 2001US-269308P.		
XX			
PA	(ELIT-) ELITRA PHARM INC.		
XX			
PI	Hasselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;		
PI	Yamamoto RT, Xu HH;		
XX			
DR	WPI; 2001-611495/70.		
DR	N-PSDB; AAS52134.		
XX			
PT	New polynucleotides for the identification and development of		
PT	antibiotics, comprise sequences of antisense nucleic acids -		
XX			
PS	Example 3; Seq ID NO 5771; 51pp; English.		
XX			
CC	The invention relates to antisense inhibitors of genes essential		
CC	prokaryotic cellular proliferation, their use in identifying the		
CC	genes, their use in the discovery of novel antibiotics, the essential		

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:41:14 ; Search time 20.5 seconds
(without alignments)
420.532 Million cell updates/sec

Title: CLAIM6
Perfect score: 1485
Sequence: 1 ADSIDINIKTGTIGSNTV.....WDRSSRYKIDWEKEMTN 293

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2.6/ptodata/1/1aa/PTUS_COMB.pep.*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	309	20.8	336	2	US-08-666-405-28
2	124.5	8.4	10182	4	US-09-134-001C-3159
3	111	7.5	461	2	US-08-392-625-24
4	111	7.5	461	2	US-08-466-961A-24
5	111	7.5	461	2	US-08-645-193B-26
6	110.5	7.4	1231	4	US-09-071-035-420
7	110.5	7.4	1265	4	US-09-071-035-418
8	108	7.3	1751	4	US-09-136-574A-44
9	106	7.1	387	4	US-09-457-302-1
10	105.5	7.1	432	1	US-08-476-008-61
11	105.5	7.1	432	1	US-08-306-063-61
12	105.5	7.1	432	1	US-08-833-485-61
13	105.5	7.1	432	4	US-09-137-440-61
14	105	7.1	1584	4	US-09-457-040B-27
15	104.5	7.0	279	4	US-09-134-001C-4878
16	103.5	7.0	440	1	US-08-061-062A-6
17	103.5	7.0	440	1	US-08-061-062A-8
18	103.5	7.0	440	3	US-08-536-150-6
19	103.5	7.0	440	3	US-08-536-150-8
20	101.5	6.8	345	4	US-09-073-898-140
21	101.5	6.8	771	3	US-08-434-000A-8
22	101.5	6.8	771	4	US-09-312-157-8
23	100.5	6.8	686	4	US-09-368-169-8
24	100	6.7	513	1	US-08-200-232-4
25	100	6.7	513	5	PCT-US95-02219-4
26	100	6.7	513	5	PCT-US95-02219A-4
27	99	6.7	537	4	US-09-110-959A-4

28	98	6.6	447	4	US-09-627-376-10	Sequence 10, Appl
29	98	6.6	464	4	US-09-426-072-2	Sequence 2, Appl
30	98	6.6	721	4	US-09-390-234-20	Sequence 20, Appl
31	98	6.6	971	2	US-08-724-354D-22	Sequence 22, Appl
32	98	6.6	971	3	US-09-270-984A-22	Sequence 22, Appl
33	98	6.6	971	3	US-09-177-431-8	Sequence 8, Appl
34	97.5	6.6	1007	4	US-08-961-083-216	Sequence 216, Appl
35	97	6.5	869	2	US-08-483-101-15	Sequence 15, Appl
36	96.5	6.5	4536	4	US-09-180-422B-27	Sequence 27, Appl
37	96	6.5	480	1	US-07-752-428E-2	Sequence 2, Appl
38	96	6.5	480	1	US-07-752-428C-2	Sequence 2, Appl
39	96	6.5	480	1	US-07-752-428C-4	Sequence 4, Appl
40	96	6.5	566	1	US-08-415-823-4	Sequence 4, Appl
41	96	6.5	566	2	US-09-086-662-4	Sequence 4, Appl
42	95	6.4	2314	4	US-09-268-347-49	Sequence 49, Appl
43	94.5	6.4	567	3	US-08-646-273-19	Sequence 19, Appl
44	94.5	6.4	708	3	US-08-646-273-23	Sequence 23, Appl
45	94.5	6.4	754	3	US-08-646-273-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-08-666-405-28
; Sequence 28, Application US/08666405
; Patent No. 5874220
; GENERAL INFORMATION:
; APPLICANT: FACH, Patrick; GUILLON,
; APPLICANT: Jean-Pierre; POPOFF, Michel
; TITLE OF INVENTION: PRIMERS FOR THE
; TITLE OF INVENTION: AMPLIFICATION OF GENES CODING FOR THE
; TITLE OF INVENTION: ENTEROTOXIN AND THE LECITHINASE OF CLOSTRIDIUM
; TITLE OF INVENTION: PEPFRINGENS AND THEIR APPLICATION TO THE
; TITLE OF INVENTION: DETECTION AND NUMERATION OF THESE BACTERIAE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/666,405
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/04292
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/172,026
; FILING DATE: 22-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: MUSERLIAN, CHARLES A
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 102.164
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:

[illegible]

ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 07-APR-1997
APPLICATION NUMBER: US/08/833,485
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-485-61

Query Match 7.1%; Score 105.5; DB 1; Length 432;
Best Local Similarity 20.1%; Pred. No. 3;
Matches 55; Conservative 46; Mismatches 89; Indels 83; Gaps 11;

QY 46 KNHKKLLVIRTKGTAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPN 105
Db 147 ENEGYPLAIRNKGKGVKVI---DGSISSQFLTALLMSAPLAENDTEIIGELVSKP 203
QY 106 SID-----TKXYXSLXYXFXGAVXGXDXGXIXGXIXAV 140
Db 204 YIDITLAMRDFGVKVENHHYQFVKVGNQSYISPNKYLVEGDASSASYFLAAGAIKGV 263
QY 141 XIXH-XLXYXQPD--FKTILESPTDKVGVKVFNNMNVNQNWPYDRSDSNPVYGNQLEM 197
Db 264 KVTGIGKNSIQGRDLFADVLE-----KMGAKI-----TWG-----EDFI 297
QY 198 KTRNGSKAANDFLDPNPKASSLSGSPDFATVITMDRKASKQOTNIDVIYERVRDQY 257
Db 298 QAEHAELNGID--MDMNH-----PDAAMTIATLFSNGETVIRNIY-----338
QY 258 LHWSTNWKGTNTKDKWDRSSERYKIDWEKEE 290
Db 339 -----NWRVKET--DRLTAMATELRKVGAEVEE 364

RESULT 13
US-09-137-440-61
; Sequence 61, Application US/09137440
; Patent No. 6248876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.

APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/137,440
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER: US 08/833,485
FILING DATE: 07-APR-1997
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-137-440-61

Query Match 7.1%; Score 105.5; DB 4; Length 432;
Best Local Similarity 20.1%; Pred. No. 3;
Matches 55; Conservative 46; Mismatches 89; Indels 83; Gaps 11;

QY 46 KNHKKLLVIRTKGTAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPN 105
Db 147 ENEGYPLAIRNKGKGVKVI---DGSISSQFLTALLMSAPLAENDTEIIGELVSKP 203
QY 106 SID-----TKXYXSLXYXFXGAVXGXDXGXIXGXIXAV 140
Db 204 YIDITLAMRDFGVKVENHHYQFVKVGNQSYISPNKYLVEGDASSASYFLAAGAIKGV 263
QY 141 XIXH-XLXYXQPD--FKTILESPTDKVGVKVFNNMNVNQNWPYDRSDSNPVYGNQLEM 197
Db 264 KVTGIGKNSIQGRDLFADVLE-----KMGAKI-----TWG-----EDFI 297
QY 198 KTRNGSKAANDFLDPNPKASSLSGSPDFATVITMDRKASKQOTNIDVIYERVRDQY 257
Db 298 QAEHAELNGID--MDMNH-----PDAAMTIATLFSNGETVIRNIY-----338
QY 258 LHWSTNWKGTNTKDKWDRSSERYKIDWEKEE 290
Db 339 -----NWRVKET--DRLTAMATELRKVGAEVEE 364

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:42:29 ; Search time 10.5 seconds

(without alignments)
1176.748 Million cell updates/sec

Title: CLAIM6

Perfect score: 1485

Sequence: 1 ADSIDINIKTGTDTIGSNTTV.....WDRSSRYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	424.5	28.6	330	10	US-09-815-242-5293
2	424.5	28.6	333	10	US-09-815-242-12121
3	404	27.2	325	10	US-09-815-242-5783
4	404	27.2	325	10	US-09-815-242-12693
5	404	27.2	325	10	US-09-815-242-12779
6	404	27.2	325	10	US-09-815-242-13070
7	403.5	27.2	327	10	US-09-815-242-5901
8	403.5	27.2	327	10	US-09-815-242-12729
9	403.5	27.2	327	10	US-09-815-242-13157
10	397	26.7	333	10	US-09-815-242-5891
11	397	26.7	338	10	US-09-815-242-13102
12	252	17.0	321	10	US-09-815-242-5771
13	252	17.0	321	10	US-09-815-242-12694
14	252	17.0	321	10	US-09-815-242-12721
15	234	15.8	312	10	US-09-815-242-5267
16	109.5	7.4	1300	10	US-09-815-242-4303
17	109.5	7.4	1300	10	US-09-815-242-10906
18	109	7.3	861	9	US-09-820-843A-109
19	106.5	7.2	771	10	US-09-818-247-4

20	106	7.1	1349	10	US-09-815-242-5898	Sequence 5898, Ap
21	106	7.1	1349	10	US-09-815-242-13137	Sequence 13137, A
22	105.5	7.1	309	10	US-09-737-178-121	Sequence 121, App
23	105.5	7.1	432	9	US-09-464-099A-61	Sequence 61, Appl
24	105.5	7.1	432	10	US-09-861-696-61	Sequence 61, Appl
25	105	7.1	978	10	US-09-815-242-5456	Sequence 5456, Ap
26	105	7.1	1001	10	US-09-815-242-12686	Sequence 12686, A
27	105	7.1	1352	9	US-09-784-554B-2	Sequence 2, Appl
28	104	7.0	791	9	US-10-055-364-41	Sequence 41, Appl
29	101.5	6.8	345	10	US-09-850-351A-140	Sequence 140, App
30	101.5	6.8	771	9	US-09-982-107-8	Sequence 8, Appl
31	100	6.7	309	10	US-09-737-178-136	Sequence 136, App
32	100	6.7	812	9	US-10-078-770-126	Sequence 126, App
33	99	6.7	1167	10	US-09-815-242-11522	Sequence 11522, A
34	98	6.6	447	12	US-10-047-676A-10	Sequence 10, Appl
35	98	6.6	941	9	US-09-992-598-353	Sequence 353, App
36	98	6.6	941	9	US-09-989-293A-353	Sequence 353, App
37	98	6.6	941	9	US-09-989-735-353	Sequence 353, App
38	98	6.6	941	9	US-09-990-444-353	Sequence 353, App
39	98	6.6	941	9	US-09-989-730-353	Sequence 353, App
40	98	6.6	941	9	US-09-990-436-353	Sequence 353, App
41	98	6.6	941	9	US-09-991-181-353	Sequence 353, App
42	98	6.6	941	9	US-09-993-687-353	Sequence 353, App
43	98	6.6	941	9	US-09-989-734-353	Sequence 353, App
44	98	6.6	941	9	US-10-028-072-464	Sequence 464, App
45	98	6.6	941	9	US-09-997-653-353	Sequence 353, App

ALIGNMENTS

RESULT 1

US-09-815-242-5293
; Sequence 5293, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5293
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5293

Query Match 28.6%; Score 424.5; DB 10; Length 330;
Best Local Similarity 30.8%; Pred. No. 4.5e-24;

Query Match 27.2%; Score 403.5; DB 10; Length 327;
Best Local Similarity 28.8%; Pred. No. 1.7e-22;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps

QY 18 TTVKTDGLVTDYDKENGSMHKRVFYSDIDNNHNKKLIVRTGTIAGQRYRVSSEGANKSG 77
||| ||| : : : : : ||| : : : : :
Dh 43 TLVKTTATSDNDKLN-TSOLLTFETKDSYKDQDTVLVAAGNINCYSKYKKDNPKDYNVSO 101

78	QY	LAWPSAFKVOQLP	NEVAQISDY	PPRNSIDTK	XXYXSLXX	YFXG	XVXG	XGXIXGXIX	137
79	QY	WYVWYVWYVWYV	WYVWYVWYV	WYVWYVWYV	WYVWYVWYV	WYVWYVWYV	WYVWYVWYV	WYVWYVWYV	138
102	Db	FYWGKYNYSV	SSNESD	AVNVVD	APKQNEE	FQVQTL	IGSYGGDI	--NISNGLSGGLN	159
103	QY	AXYXIXHXLXX	YXQDPDK	TILESPD	--KKYCWK	VIFFNMV	NQNWGP	YDRDSWHPVYNOLF	196
104	QY	WYVWYVWYVWYV	WYVWYVWYV	WYVWYVWYV	WYVWYVWYV	WYVWYVWYV	WYVWYVWYV	WYVWYVWYV	197
160	Db	GSKSFSET	INYKQESYRT	IDRKNHKS	IGWGEAH	KIMNNG	GWPGYR	SDSYDPTYGNEFL	219
197	QY	MKTRNG	SMKAAFL	DPNKA	SSLLSG	FSDFAT	VTIMDR	KASKQOTNTDVIYERVDDY	256
220	Db	LGGRQSS	NAGQNF	LTHOM	PLLARG	NFPN	PEFISV	LSHKQNTK--KSKTKYIVQREMDRY	278
257	QY	QLHWT	STNKGNT	KTDKW	TRDSR	SERYKIDW	E		287
279	Db	TNOWNEL	LHWYGN	NYKNONT	VTFTST	VEYDWO			309

```

/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ PROKARYOTES
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-05-26

```

```
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12729
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12729

Query Match 27.2%; Score 403.5; DB 10; Length 327;
Best Local Similarity 28.8%; Pred. No. 1.7e-22;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;

QY 18 TTVKGTDLVTDKENGHMKVFYFIDDKNNHKKLLVIRTKGTAGQYRVYSBEGANKSG 77
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
43 TLYKTTATSDNDKLN-ISOILTENFIKDSYDKDTLVKKAAGNINSYKPKPNKDYNYSQ 101

QY 78 LWPFAFKVOLQLPDNEVAQISDYPRNSIDTKYXSLXLYXFXGXVXGXDXGXIXGXIX 137
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 102 FYWGGKYNVSSESNDVAVNVDYAPKNQNEEFQVOOTLGYSYGGDI--NISNLSGGLN 159

QY 138 AXVXIXHLXAYXOPDFKTIESTPD-KKVGKVIFFNNVNNQNGPYDRDSWNPVYGQOLF 196
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 160 GSKFSFSTINYKQESYRTTIDRKNHKSIGWVEAHKIMNNGWGPYGRDSYDPTYGNELF 219

QY 197 MKTRNGSMKAADNPLDPNKASSLLSGFSPDFATVITMDRKASKQOQTNIIVYERVRDDY 256
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 220 LGGRQSSNAGQFLPHQMPLLARGNFPFISVLSHKQNDYK-KSKIKVTYQREMDRY 278

QY 257 QLHWTSTNWKTNTKDKWTRSSERYKIDWE 287
| | | | | : : : : : | | | | | : : : : : |
Db 279 TNQWNRLLHWVGNVYKQNTVTFTSTYEVDWQ 309

RESULT 9
US-09-815-242-13157
; Sequence 13157, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12729
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12729

Query Match 26.7%; Score 397; DB 10; Length 323;
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13157
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13157

Query Match 27.2%; Score 403.5; DB 10; Length 327;
Best Local Similarity 28.8%; Pred. No. 1.7e-22;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;

QY 18 TTVKGTDLVTDKENGHMKVFYFIDDKNNHKKLLVIRTKGTAGQYRVYSBEGANKSG 77
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
43 TLYKTTATSDNDKLN-ISOILTENFIKDSYDKDTLVKKAAGNINSYKPKPNKDYNYSQ 101

QY 78 LWPFAFKVOLQLPDNEVAQISDYPRNSIDTKYXSLXLYXFXGXVXGXDXGXIXGXIX 137
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 102 FYWGGKYNVSSESNDVAVNVDYAPKNQNEEFQVOOTLGYSYGGDI--NISNLSGGLN 159

QY 138 AXVXIXHLXAYXOPDFKTIESTPD-KKVGKVIFFNNVNNQNGPYDRDSWNPVYGQOLF 196
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 160 GSKFSFSTINYKQESYRTTIDRKNHKSIGWVEAHKIMNNGWGPYGRDSYDPTYGNELF 219

QY 197 MKTRNGSMKAADNPLDPNKASSLLSGFSPDFATVITMDRKASKQOQTNIIVYERVRDDY 256
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 220 LGGRQSSNAGQFLPHQMPLLARGNFPFISVLSHKQNDYK-KSKIKVTYQREMDRY 278

QY 257 QLHWTSTNWKTNTKDKWTRSSERYKIDWE 287
| | | | | : : : : : | | | | | : : : : : |
Db 279 TNQWNRLLHWVGNVYKQNTVTFTSTYEVDWQ 309

RESULT 10
US-09-815-242-5891
; Sequence 5891, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5891
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5891

Query Match 26.7%; Score 397; DB 10; Length 323;
```

```
Best Local Similarity 28.5%; Pred. No. 5.1e-22;
Matches 81; Conservative 83; Mismatches 108; Indels 12; Gaps 7;

QY 13 DIGSNTTVKTGLVTDKENGHHKVFYFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEG 72
Db 36 NLDGDFKMTYTRATTSDSQKNITQSLOFNFLTEPNYDKETVFIKAGTTIGSLRILDPNG 95
QY 73 ANKSGLAWPSAFKVLQ-LPDNEVAQISDYPRNSIDTKXYXSLXYFXGXVXGXDXGX 131
Db 96 YWNSTLRWPGSYSVSIQNVDNNNTNTDFAPKNQDESREVKYTYGKTGGD-FSINRGG 154
QY 132 IXGXIXAXVIXHXLXYXQDPFKTTILESTDKK-VGWKV---IFNNMVNQNGPYDRDSW 187
Db 155 LTGNITKESNYSETISYQOQPSYRTLLDQSTSHKGVGWVKEAHLINNMGHDHTRQLTNDSD 214
QY 188 NPVYGOLFMTKTRNGSMKAADNFDLPNKASSLLSSGFSDFATVITMDRKASKQOQTNDIV 247
Db 215 NRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTVSGFNPEFLAVMSHDKK-DKGRKSQFVV 272
QY 248 IYERVDDYQLHWTSTN----WKGNTKDKWTDTRSERYKIDWE 287
Db 273 HYKRSMDFEKIDNRHGFNGWGSNGENHVDKKEKLSALYEVDWK 316

RESULT 11
US-09-815-242-13102
; Sequence 13102, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13102
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13102

Query Match 26.7%; Score 397; DB 10; Length 338;
Best Local Similarity 28.5%; Pred. No. 5.5e-22;
Matches 81; Conservative 83; Mismatches 108; Indels 12; Gaps 7;

QY 13 DIGSNTTVKTGLVTDKENGHHKVFYFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEG 72
Db 42 NLDGDFKMTYTRATTSDSQKNITQSLOFNFLTEPNYDKETVFIKAGTTIGSLRILDPNG 101
QY 73 ANKSGLAWPSAFKVLQ-LPDNEVAQISDYPRNSIDTKXYXSLXYFXGXVXGXDXGX 131

Best Local Similarity 28.5%; Pred. No. 5.1e-22;
Matches 81; Conservative 83; Mismatches 108; Indels 12; Gaps 7;

QY 13 DIGSNTTV--KTGDLVTDKENGHHKVFYFIDDKNNHKKLLVIRTKGTIAGQYRVYSE 70
Db 47 DIGOGAEIIRKTDITS--KRLAITQNIQDFVVDKYNKDALVYKMGGFSSR-TTYSYD 103
QY 71 --EGANKSGLAWPSAFKVLQ-LPDNEVAQISDYPRNSIDTKXYXSLXYFXGXVXGXDX 128
Db 104 LKPYIKRMIWPFQYINSLTKDSNV-DLINYLPPKNKIDSADYSQKLYGNIIGNFQFAP 162
QY 129 XGXIXGXIXAXVIXHXLXYXQDPFKTTILESTDKKVGWVKEAHLINNMGHDHTRQLTNDSD 188
Db 163 --STGG--SGSFNTSKTISYQKNYKIVTEVESQNSKGVKGVKANSFVTPN-----GQV 211
QY 189 PVYGNOLFMTKTRNGSMKAADNFDLPNKASSLLSSGFSDFATVITMDR-KASKOQTNDIV 247
```

```
Db 102 YWNSTLRWPGSYSVSIQNVDNNNTNTDFAPKNQDESREVKYTYGKTGGD-FSINRGG 160
QY 132 IXGXIXAXVIXHXLXYXQDPFKTTILESTDKK-VGWKV---IFNNMVNQNGPYDRDSW 187
Db 161 LTGNITKESNYSETISYQOQPSYRTLLDQSTSHKGVGWVKEAHLINNMGHDHTRQLTNDSD 220
QY 188 NPVYGOLFMTKTRNGSMKAADNFDLPNKASSLLSSGFSDFATVITMDRKASKOQTNDIV 247
Db 221 NRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTVSGFNPEFLAVMSHDKK-DKGRKSQFVV 278
QY 248 IYERVDDYQLHWTSTN----WKGNTKDKWTDTRSERYKIDWE 287
Db 279 HYKRSMDFEKIDNRHGFNGWGSNGENHVDKKEKLSALYEVDWK 322

RESULT 12
US-09-815-242-5771
; Sequence 5771, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5771
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5771

Query Match 17.0%; Score 252; DB 10; Length 321;
Best Local Similarity 24.9%; Pred. No. 4.3e-11;
Matches 71; Conservative 71; Mismatches 119; Indels 24; Gaps 12;

QY 13 DIGSNTTV--KTGDLVTDKENGHHKVFYFIDDKNNHKKLLVIRTKGTIAGQYRVYSE 70
Db 47 DIGOGAEIIRKTDITS--KRLAITQNIQDFVVDKYNKDALVYKMGGFSSR-TTYSYD 103
QY 71 --EGANKSGLAWPSAFKVLQ-LPDNEVAQISDYPRNSIDTKXYXSLXYFXGXVXGXDX 128
Db 104 LKPYIKRMIWPFQYINSLTKDSNV-DLINYLPPKNKIDSADYSQKLYGNIIGNFQFAP 162
QY 129 XGXIXGXIXAXVIXHXLXYXQDPFKTTILESTDKKVGWVKEAHLINNMGHDHTRQLTNDSD 188
Db 163 --STGG--SGSFNTSKTISYQKNYKIVTEVESQNSKGVKGVKANSFVTPN-----GQV 211
QY 189 PVYGNOLFMTKTRNGSMKAADNFDLPNKASSLLSSGFSDFATVITMDR-KASKOQTNDIV 247
```

```

; sequence 2607; application 05/05010244
; Patent No. US2000061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:40:49 ; search time 27.5 seconds
(without alignments)
1024.269 Million cell updates/sec

Title: CLAIM6

Perfect score: 1485

Sequence: 1 ADSIDINIKTGTIGSNTTV.....WTDRSSERYKIDWEKEEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	1485	100.0	319	2 S69209	alpha-toxin precursor
2	1481	99.7	319	2 F89887	Alpha-Hemolysin precursor
3	409	27.5	325	2 B42334	leucocidin R, comp
4	404	27.2	300	2 B49238	gamma-hemolysin ga
5	404	27.2	325	2 F90043	gamma-hemolysin co
6	402.5	27.1	327	2 B89968	leukotoxin, LukD I
7	399	26.9	338	2 C89991	hypothetical prote
8	396	26.7	325	2 T00161	leucocidin chain F
9	394.5	26.6	326	2 S68224	synergohymenotropl
10	393	26.5	325	2 S32212	leucocidin chain F
11	390	26.3	323	2 JQ1530	leukocidin chain F
12	390	26.3	323	2 JQ6027	leukocidin chain F
13	383	25.8	322	2 JC5469	Panton-Valentine l
14	309	20.8	336	2 I40826	beta-toxin - Clost
15	257	17.3	308	2 JC3468	leucocidin chain I
16	255.5	17.2	351	2 D89991	hypothetical prote
17	252	17.0	309	2 D90043	gamma-hemolysin ch
18	251	16.9	309	2 JC4282	gamma-hemolysin II
19	245	16.5	321	2 S49271	hlga-like protein
20	244	16.4	286	2 C49238	gamma-hemolysin co
21	244	16.4	315	2 PC4078	hlgc-like protein
22	244	16.4	315	2 E90043	gamma-hemolysin co
23	241.5	16.3	311	2 C89968	leukotoxin LukE I1
24	241	16.2	315	2 A49234	leucocidin R S com
25	238	16.0	315	2 JN0626	leucocidin chain S
26	237	16.0	312	2 T00160	leucocidin chain S
27	233	15.7	312	2 S32211	leucocidin chain S
28	201.5	13.6	310	2 S68225	synergohymenotropl
29	121	8.1	1562	2 S53069	probable membrane

30	115.5	7.8	1233	1 G71612	novel serine/threo
31	115.5	7.8	1711	2 C71625	variant-specific s
32	114	7.7	1386	2 AC1533	surface protein (L
33	113.5	7.6	550	2 G86667	hypothetical prote
34	113.5	7.6	796	2 A90541	hypothetical prote
35	113.5	7.6	1787	2 T20160	hypothetical prote
36	113	7.6	417	2 G64417	hypothetical prote
37	113	7.6	435	2 G86907	D-alanyl-D-alanine
38	113	7.6	588	2 B71802	penicillin-binding
39	112.5	7.6	453	2 H70352	mannose-6-phosphat
40	112	7.5	976	2 A42466	alpha-amylase (EC
41	111.5	7.5	230	2 JE0083	luza chaperon - Ph
42	111.5	7.5	569	2 A45624	trophozoite cystel
43	111.5	7.5	928	2 S54699	hemopexin-heme com
44	111	7.5	461	2 S23420	probable subtilisi
45	110.5	7.4	437	2 T03553	probable maltose-b

ALIGNMENTS

RESULT 1

S69209

alpha-toxin precursor - Staphylococcus aureus

C:Species: Staphylococcus aureus

C>Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000

C:Accession: S69209

R:Hedengrahn, G. submitted to the EMBL Data Library, October 1992

A:Reference number: S69209

A:Accession: S69209

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-319 <HED>

A:Cross-references: EMBL:X01645; NID:g46763; PIDN:CAA25801.1; PID:g46765

C:Superfamily: leukocidin

C:Keywords: toxin

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-319/Product: alpha-toxin #status predicted <MAT>

Query Match 100.0%; Score 1485; DB 2; Length 319;
Best Local Similarity 93.2%; Pred. No. 6.8e-89;
Matches 273; Conservative 20; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADSIDINIKTGTIGSNTTVKGTGDLVYDKENGMMHKVFYSFIDDKNNHKKLLVITKGT 60

Db 27 ADSIDINIKTGTIGSNTTVKGTGDLVYDKENGMMHKVFYSFIDDKNNHKKLLVITKGT 86

QY 61 IAGQYRVYSEEGANKSGLAWPSAFKVLQQLPDNEVAQISDYYPNSIDTKXXYXSLXYXF 120

Db 87 IAGQYRVYSEEGANKSGLAWPSAFKVLQQLPDNEVAQISDYYPNSIDTKXXYXSLXYXF 146

QY 121 XGVXGXDXGXIXGXIXAXYXIXHXLXIXQDPFKTILESTDKKVGKVIFFNNVQNWG 180

Db 147 NGVNTGDDTGKIGGLIGANVSIGHTLKYVQDPFKTILESTDKKVGKVIFFNNVQNWG 206

QY 181 PYDRDSNPNVYGNQLPMKPFNGSKAADNFDLPNKASLLSSGFSDFATVITMDRKASK 240

Db 207 PYDRDSNPNVYGNQLPMKPFNGSKAADNFDLPNKASLLSSGFSDFATVITMDRKASK 266

QY 241 QQTNDIVYVERVDDYQLHWTSTNWKGTNTKDKWTRSSERYKIDWEKEEMTN 293

Db 267 QQTNDIVYVERVDDYQLHWTSTNWKGTNTKDKWTRSSERYKIDWEKEEMTN 319

RESULT 2

F89887

Alpha-Hemolysin precursor [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 01-Mar-2002

C:Accession: F89887

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Sekimizu, Y. ma, A.; Mizutani-Ul, Y.; Kobayashi, N.; Savano, R.; Inoue, R.; Kaito, C.;

leukocidin chain F precursor - Staphylococcus aureus phage phi pVL
T00161
N;Alternate names: protein 28
C;Species: Staphylococcus aureus phage phi pVL
C;Date: 23-Apr-1999 #sequence revision 23-Apr-1999 #text change 21-Jul-2000

QY 255 DIQLHWISTNNKGTINAKRWIDRRSSERIKIDWEK 288
: ! : ! ! ! ! : : ! : ! ! ! :
Db 277 EYTNYNWNGFHWMGTYNQNONNATFSEYFIEDWQ 310

```

RESULT 10
S32212
leucocidin chain F - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Mar-2000
C:Accession: S32212
R:Prevost, G.; Supersac, G.; Piemont, Y.
submitted to the EMBL data Library, March 1993
A:Description: Staphylococcus aureus encodes two types of synergohymenotro
A:Reference number: S32211
A:Accession: S32212
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <PRE>
A:Cross-references: EMBL:X72700; NID:g551668; PIDN:CAA51252.1; PID:g388292
C:Genetics:
A:Gene: lukF
C:Superfamily: leucocidin
C:Keywords: toxin

```

[illegible]

RESULT 11
JQ1530
leukocidin chain F precursor - Staphylococcus aureus
N;Alternate names: gamma-hemolysin chain H gamma I
C;Species: Staphylococcus aureus
C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-2001
C;Accession: JQ1530; PO0336; S32418

[illegible]

QY 106 SIDTKYXSLXVFXGXVGDGXGXIXXAXVIXHXLXVXQDPDFKTIESTPTD-KK 164
Db 130 QNEEFQVQTLGYSYGDDI--NIINGLTGLGNGSKSFSETINKQESYRTTIDRKTNHS 187
QY 165 VGNKVFNNMNVNMGPGYDRDSDNPNVYGNQLFPMKTRNGSKAADNFDLPNKASSLLSSGF 224
Db 188 IGVGVFAHKIMNGWPGYDRDSDSLYGNEFLFGGQSSNANQNQLPHTQPIIARGNF 247
QY 225 SPDFATVITMDRKASKQOTNIDVIYVRDDYQLHWTSTNWKGTNFKDKWTDSSRYKI 284
Db 248 NPEFISVLSHKQDKVK-KSKIKVYQREMDRYENFNHNLHWGYNKKNQKRATHTSIYEL 306
QY 285 DWK 288
Db 307 DWK 310
RESULT 14
beta-toxin - Clostridium perfringens
C:Species: Clostridium perfringens
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 17-Nov-2000
C:Accession: I40826; I40856; S51649
R:Hunter, J.E.C.; Bronn, J.E.; Oyston, P.C.F.; Jakura, J.; Titball, R.W.
Infect. Immun. 61, 3958-3965, 1993
A:Title: Molecular genetic analysis of the beta-toxin of Clostridium perfringens re-
A:Reference number: I40826; MUID:93366460; PMID:8359918
A:Accession: I40826
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-336 <RES>
A:Cross-references: GB:L13198; NID:9410018; PIDN:AAA23284.1; PID:9410019
R:Steinporstottir, V.; Frithriksdottir, V.; Gunnarsson, E.; Andresson, O.S.
FEMS Microbiol. Lett. 130, 273-278, 1995
A:Title: Expression and purification of Clostridium perfringens beta-toxin glutathione S
A:Reference number: I40856; MUID:93377614; PMID:7649450
A:Accession: I40856
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-170, 'K', 172-336 <RE2>
A:Cross-references: EMBL:X83275; NID:9602993; PIDN:CAA58246.1; PID:9602994
C:Genetics:
A:Gene: cfp
A:Superfamily: leukocidin
C:Keywords: toxin

Query Match 20.8%; Score 309; DB 2; Length 336;
Best Local Similarity 25.3%; Pred. No. 2.5e-12;
Matches 80; Conservative 75; Mismatches 115; Indels 46; Gaps 11;
QY 13 DIGSNTTV---KTGD-----LVTYD-----KENGMMKKVYFSDIDKNNHKK 51
Db 29 DIGKTTTTRNKTSDGYTIITQNDQIISYQSDVSSSKNEDGFTASIDARFIDDKYSSE 88
QY 52 LLVIRTKGTIAG-----QYRVSEEGANKSGLWPSAFKVO-IQLPDNEVAQISDYYP 103
Db 89 TFLNLTGFMSSKEDVDIKKYNLH--DVTNSTAINFPVYRYSISILNESINENVKIVDSIP 146
QY 104 RNSIDTKYXSLXVFXGXVGDGXGXIXXAXVIXHXLXVXQDPDFKTIESTPTD 163
Db 147 KNTISQKVTNMGKIGGSIE-IEENPKASIESEYAESSTIEYQDPDFSTIQDHTS 205
QY 164 KYGKVFNNMNVNMGPGYDRDSDNPNVYGNQLFPMKTRNGSKAADNFDLPNKASSLLSSG 223
Db 206 KASWDTKFTETTR---GYNLKNPNPVYGNEMFYGRYTNVATENIIPDYQMSKLIITGG 262
QY 224 FSPDFATVITMDRKASKQOTNIDVIYVRDDYQLHWTSTNWKGT-----NTRDKWTD 276
Db 263 LNPNNSVVLTAPN--GTEESIIKVMERNCYILNNGANNVGVQVYSRLAFDTPN--VD 318
QY 277 RSSERYKIDWEKEMT 292
Db 319 SHIFTKINLWTHKVT 334

RESULT 15

JC5468

leukocidin chain lukM precursor - Staphylococcus aureus

C:Species: Staphylococcus aureus

C>Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 20-Jun-2000

C:Accession: JC5468; PC4338; S68460; S68542

R:Kaneko, J.; Muramoto, K.; Kamio, Y.

Biosci. Biotechnol. Biochem. 61, 541-544, 1997

A:Title: Gene of LukF-PV-like component of Pantan-Valentine leukocidin in Staphylococ

A:Reference number: JC5468; MUID:97249691; PMID:9095557

A:Accession: JC5468

A:Molecule type: DNA

A:Residues: 1-308 <KAN>

A:Cross-references: DDBJ:D83951; NID:gl230553; PIDN:BAA12147.1; PID:gl230554

A:Experimental source: strain P83

A:Accession: PC4338

A:Molecule type: protein

A:Residues: 29-75;101-130;263-283 <KA2>

R:Choorit, W.; Kaneko, J.; Muramoto, K.; Kamio, Y.

FEBS Lett. 357, 260-264, 1995

A:Title: Existence of a new protein component with the same function as the LukF comp

A:Reference number: S68460; MUID:95137101; PMID:7835424

A:Accession: S68460

A:Molecule type: DNA

A:Residues: 1-2, 'N', 4-288, 'L', 290-299, 'S', 301, 'L', 303-304, 'R', 306-308 <CHO>

A:Cross-references: EMBL:D42144; NID:9577648; PIDN:BAA07715.1; PID:9577649

A:Experimental source: strain P83 (ATCC 31890)

A:Accession: S68542

A:Molecule type: protein

A:Residues: 29-75;101-130;263-282 <CHW>

C:Genetics:

A:Gene: lukM

C:Function:

C:Description: causes cytotoxic changes in polymorphonuclear leukocytes in human and

C:Superfamily: leukocidin

C:Keywords: toxin

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-308/Product: leukocidin chain lukM #status experimental <MAT>

Query Match 17.3%; Score 257; DB 2; Length 308;

Best Local Similarity 24.1%; Pred. No. 5.2e-09;

Matches 70; Conservative 74; Mismatches 121; Indels 26; Gaps 12;

QY 9 TGTDTIGSNTTV--KTGDLVTYDKENGMMKKVYFSDIDKNNHKKLLVIRTKGTIAGQ-- 64
Db 30 TNAEDIGDDAEVIKRTEDVSS--RRWGVTONVQDFVKKYKNDALIKMGQFINSRRT 87
QY 65 YRVYSEEGANKSGLWPSAFKVOQLPDNEVAQISDYYPNRSIDTKYXSLXVFXGXV 124
Db 88 FNDVKQNRANKR--WVWPFQYNIGLTSKQNTSLI--NYPKNNKIEVDVQTLGYNIGKF 145
QY 125 XGXDXGXIXGXIXAXVIXHXLXVXQDPDFKTIESTPTDKKVGKVFNNMV--NONWGPY 182
Db 146 QSVF--SIGG--NGSFNYSKISYQSVSEVEQSSKTIKGVKANSFVIAGHRWSAY 201
QY 183 DRDSNPNVYGNQLFPMKTRNGSKAADNFDLPNKASSLLSSGSPDFATVITMDRKASKQ 242
Db 202 D-----ELLTRNTTRGNARDYFDDNPLPLTSGFNFSFIATVSHE--KDSGDT 251
QY 243 TNIDVIYVRD--DYQLHWTSTNWKGTNTRDKWTD--SERIKIDWEKEEM 291
Db 252 SEFETIGRNDVIYATYLPKGLYPERKHNEFVNRNPFVYKVEYVNWKTIEI 302

Search completed: March 4, 2003, 10:45:09

Job time : 29.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:39:33 ; Search time 17 Seconds
(without alignments)
714.857 Million cell updates/sec

Title: CLAIM6

Perfect score: 1485

Sequence: 1 ADSDNINIKGTGTDIGSNTV.....WDRSSERYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1485	100.0	319	1 HLA_STAAU	P09616 staphylococ
2	404	27.2	325	1 HLBG_STAAU	Q07226 staphylococ
3	390	26.3	323	1 LUKE_STAAU	P31715 staphylococ
4	252	17.0	309	1 HLGA_STAAU	P31714 staphylococ
5	244	16.4	315	1 HLCG_STAAU	Q07227 staphylococ
6	238	16.0	315	1 LUKS_STAAU	P31716 staphylococ
7	121	8.1	1562	1 YM81_YEAST	Q04781 saccharomyc
8	113.5	7.6	1787	1 CHD3_CAEEL	Q22516 caenorhabdi
9	113	7.6	417	1 Y943_METJA	Q58353 methanococc
10	112	7.5	976	1 AMY_BUTFI	P30269 butyrivibri
11	111.5	7.5	569	1 CYSP_PLAFA	P25805 plasmodium
12	111.5	7.5	928	1 HXAZ_HAEIN	P45354 haemophilus
13	111	7.5	461	1 EPIP_STAEP	P30199 staphylococ
14	108.5	7.3	1786	1 YCF1_ARATH	P56785 arabidopsis
15	107.5	7.2	420	1 PIGR_MYCLE	Q69473 mycobacteri
16	106.5	7.2	771	1 PIGR_MOUSE	O70570 mus musculu
17	106.5	7.2	1066	1 HXSA_STRPN	O54873 streptococc
18	106	7.1	387	1 M5OX_BACSN	P23342 bacillus sp
19	105.5	7.1	380	1 GUG2_YEAST	P47011 saccharomyc
20	105	7.1	464	1 MRJ4_APIME	Q17061 apis mellif
21	105	7.1	859	1 RPA2_METJA	Q58446 methanococc
22	105	7.1	1584	1 KYK1_DICDI	P18160 dictyosteli
23	105	7.1	1655	1 OMPB_RICCN	Q9Kka3 r outer mem
24	104	7.0	1656	1 OMPB_RICJA	Q06653 r outer mem
25	103.5	7.0	440	1 RYGA_ASPAC	Q00001 aspergillus
26	103	6.9	1212	1 YK64_CAEEL	P34343 caenorhabdi
27	103	6.9	1698	1 4L_DROME	Q9v8r9 drosophila
28	102	6.9	784	1 OSTA_ECOLI	P31554 escherichia
29	101.5	6.8	1160	1 TFC3_YEAST	P34111 saccharomyc
30	101	6.8	467	1 RGSB_HUMAN	Q94810 homo sapien
31	101	6.8	679	1 RRP1_DROME	P27864 drosophila
32	101	6.8	1007	1 RGAI_YEAST	P39083 saccharomyc
33	101	6.8	1310	1 VAC3_HELPY	Q48253 helicobacte

34	101	6.8	1365	1	GTFS_STRDO	P29336 streptococc
35	100.5	6.8	441	1	YY07_METJA	O60306 methanococc
36	100.5	6.8	556	1	Y281_MYCCE	P47523 mycoplasma
37	100.5	6.8	705	1	PPCE_FLAME	P27028 flavobacter
38	100	6.7	649	1	GUND_CLOTM	P04954 clostridium
39	99.5	6.7	432	1	AROA_HAEIN	Q03421 haemophilus
40	99	6.7	441	1	YKF9_YEAST	P35728 saccharomyc
41	99	6.7	510	1	INOL_NICPA	Q9ssv4 nicotiana p
42	99	6.7	1018	1	FNBA_STAAU	P14738 staphylococ
43	99	6.7	1592	1	GTFF2_STRDO	P27470 streptococc
44	99	6.7	1597	1	GTFF1_STRDO	P11001 streptococc
45	99	6.7	1783	1	Y468_MYCCE	Q49460 mycoplasma

ALIGNMENTS

RESULT 1
HLA_STAAU STANDARD; PRT; 319 AA.
AC P09616;
DT 01-MAR-1989 (Rel. 10, Created)
DI 01-DEC-1992 (Rel. 24, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-hemolysin precursor (Alpha-toxin) (Alpha-HL).
GN HLY OR HLA.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-44.
RC STRAIN=Wood 46;
RX MEDLINE=85053471; PubMed=6500704;
RA Gray G.S., Kehoe M.;
RT "Primary sequence of the alpha-toxin gene from Staphylococcus aureus
wood 46.";
RL Infect. Immun. 46:615-618(1984).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RC STRAIN=Wood 46;
RA Hedengrahn G.;
RN [3]
RP SEQUENCE OF 27-319 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Wood 46;
RX MEDLINE=92268149; PubMed=1587866;
RA Walker B., Krishnasastri M., Zorn L., Kasianowicz J., Bayley H.;
RT "Functional expression of the alpha-hemolysin of Staphylococcus
aureus in intact Escherichia coli and in cell lysates. Deletion of
five C-terminal amino acids selectively impairs hemolytic
activity.";
RL J. Biol. Chem. 267:10902-10909(1992).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RC STRAIN=Wood 46 / APCC 10832;
RX MEDLINE=97102581; PubMed=8943190;
RA Song L., Hobough M.R., Shustak C., Cheley S., Bayley H., Gouaux J.E.;
RT "Structure of staphylococcal alpha-hemolysin, a heptameric
transmembrane pore.";
RL Science 274:1859-1866(1996).
RN [5]
RP MUTAGENESIS.
RC MEDLINE=930161135; PubMed=1400487;
RA Walker B., Krishnasastri M., Zorn L., Bayley H.;
RT "Assembly of the oligomeric membrane pore formed by staphylococcal
alpha-hemolysin examined by truncation mutagenesis.";
RL J. Biol. Chem. 267:21782-21786(1992).
RN [6]
RP MUTAGENESIS OF HISTIDINE RESIDUES.
RC MEDLINE=94222552; PubMed=8168947;
RA Menzies B.E., Kernodle D.S.;
RT "Site-directed mutagenesis of the alpha-toxin gene of Staphylococcus
aureus: role of histidines in toxin activity in vitro and in a murine
RT

Strain	Genotype	Hemolytic Activity
285	H ⁻ L: 46% of normal hemolytic activity	285
285	H ⁻ L: 46% of normal hemolytic activity	285

[illegible]

RESULT 2

325 AA.
PRT;
STANDARD.

ID	HIGH-STRAIN STANDARD;
AC	00726;
AT	16-OCT-2001 (Rel. 40, Created)
DC	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2002 (Rel. 41, Last annotation update)
DT	15-JUN-2002 (Rel. 41, Last annotation update) (H-gamma-I).
DE	GAMMA-Hemolysin component B precursor (H-gamma-1)
DN	Hemolysin
GE	HGB OR SAZ29 OR MW2344.
GN	HLG OR HAZ29 OR MW2344.
OS	Staphylococcus aureus (strain N315), and
OS	Staphylococcus aureus (strain MW2), and
OS	Staphylococcus aureus; Bacillales; Staphylococcus.
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX	NCB1_TaxID=158879, 196620, 1280;
[1]	SEQUENCE FROM N.A.
RN	STRAIN=N315;
RC	MEDLINE=21311952; PubMed=11418146; Baba T., Yuzawa H., Kobayashi I.,
RX	Kuroda M., Ohta T., Uchiyama I., Nagai Y., Lian J.-Q., Ito T.,
RA	Cui L., Oguchi A., Aoki K.-I., Maruyama A., Murakami H., Hosoyama A.,
RA	Kanamori M., Matsumaru H., Sawano T., Inoue R.-Y., Kaito C.,
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-Y., Kaito C.,
RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Shiba T.,
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus
RT	aureus". 257.1225-1240(2001).

[2]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=HW2;
RE MEDLINE=22040471; PubMed=12044378;
RX BABA T., TAKEUCHI F., KURODA M., YUZAWA H., AOKI K.-I., OGUCHI A.,
RA BABA T., TAKEUCHI F., KURODA H., KURODA T., NAIMI T.,
RA Nagai Y., Iwana N., Asano K.,
RA Yanamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA";
RT acquired MRSA";
RT acquired MRSA";

[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=93138806; PubMed=8423103;
RX Cooney J.C., Kienle Z., Foster T.J., O'Toole P.W.;
RA "The gamma-hemolysin locus of *Staphylococcus aureus* comprises three
RT linked genes, two of which are identical to the genes for the F and S
RT components of leukocidin";
RT

model.";
[7] Infect. Immun. 62:1843-1847(1994).
[7]
MUTAGENESIS OF HISTIDINE RESIDUES.
MEDLINE=94245329; PubMed=8188346;
Jursch R., Hildebrand A., Hobom G., Tranum-Jensen J., Ward R.,
Kehoe M., Bhakdi S.; near the N-terminus of staphylococcal alpha-toxin
as reporters of regions that are critical for oligomerization and
pore formation.";
Infect. Immun. 62:2249-2256(1994).
[8]
MUTAGENESIS.
MEDLINE=96032742; PubMed=7559447;
Walker B., Bayley H.;
"Key residues for membrane binding, oligomerization, and pore forming
activity of staphylococcal alpha-hemolysin identified by cysteine
scanning mutagenesis and targeted chemical modification.";
J. Biol. Chem. 270:23065-23071(1995).
-I- FUNCTION: ALPHA-TOXIN BINDS TO THE MEMBRANE OF EUKARYOTIC CELLS
RESULTING IN AN EVENTUAL OSMOTIC LYSIS. HEPTAMER OLIGOMERIZATION
AND PORE FORMATION IS REQUIRED FOR LYtic ACTIVITY.
-II- SUBUNIT: SELF-ASSEMBLES TO FORM FIRST, A NONLYTIC OLIGOMERIC
INTERMEDIATE, AND THEN, A MUSHROOM-SHAPED HOMOEPTAMER IN DIAMETER.
OF 100 ANGSTROMS IN LENGTH AND UP TO 100 ANGSTROMS IN DIAMETER.
-III- SUBCELLULAR LOCATION: SECRETED AS A MONOMER. AFTER OLIGOMERIZATION
THE COMPLEX IS TRANSLOCATED ACROSS THE
BILAYER, PROBABLY VIA THE GLY-RICH DOMAIN OF EACH STRAND.
-IV- DOMAIN: THE MUSHROOM-SHAPED HEPTAMER IS COMPOSED OF A CAP DOMAIN
(COMPRISES 7 BETA SANDWICHES AND THE AMINO LATCHES OF EACH
PROTOMER), 7 RIM REGIONS WHOSE PROTRUDING STRANDS MAY INTERACT
WITH THE MEMBRANE BILAYER, AND THE STEM DOMAIN (52 ANGSTROMS IN
LENGTH, 26 ANGSTROMS IN DIAMETER) WHICH FORMS THE TRANSMEMBRANE
PORE.
-V- SIMILARITY: BELONGS TO THE AEROOLYSIN FAMILY.
THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation at
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; X01645; CRA25801.1; --
DR EMBL; W90536; AAA26598.1; --
PDB; 7AHL; 14-JAN-98.
InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemalysn.pore.
DR InterPro; IPR003963; Staph_bln_n_tkn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TRIFRAMS; TIGR01002; hlyii; 1.
DR PROSITE; PS00274; AEROOLYSIN; 1.
KW Hemolysis; Toxin; Signal; 3D-structure.
FT SIGNAL
FT CHAIN 1 .. 26
FT FT 27 319
FT CHAIN 145 169
FT DOMAIN 159 160
FT SITE
FT SITE 161 162
FT SITE 164 165
FT SITE 165 166
FT MUTAGEN 61 61
FT MUTAGEN 61 61
FT MUTAGEN 74 74
FT MUTAGEN 170 170

ALPHA-HEMOLYSIN.

GLY-RICH.

CLEAVAGE OF MONOMERS BY PROTEINASE K

(MINOR SITE).

CLEAVAGE OF MONOMERS BY PROTEINASE K

(MAJOR SITE).

CLEAVAGE OF MONOMERS BY PROTEINASE K

(MINOR SITE).

CLEAVAGE OF MONOMERS BY PROTEINASE K

(MAJOR SITE)

H->L: NO OLIGOMERIZATION NOR HEMOLYTIC ACTIVITY.

H->R: NO OLIGOMERIZATION NOR HEMOLYTIC ACTIVITY.

H->L: 7% OF NORMAL HEMOLYTIC ACTIVITY.

H->L: 16% OF NORMAL HEMOLYTIC ACTIVITY.


```

RL Infect. Immun. 61:768-771(1993).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89310396; PubMed=3075655;
RA Cooney J.C., Mulvey M., Arbutnot J., Foster T.;
RT "Molecular cloning and genetic analysis of the determinant for gamma-
RT lysin, a two-component toxin of Staphylococcus aureus.";
RT J. Gen. Microbiol. 134:2179-2188(1988).
CC -!- FUNCTION: GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.
CC -!- SUBUNIT: CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I AND -II).
CC H-GAMMA-I IS ALSO THE F SUBUNIT OF LEUKOCIDIN.
CC -!- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP003137; BAB43511.1; -.
DR EMBL; AP004830; BAB96209.1; -.
DR EMBL; L01055; AAB26639.1; -.
DR HSP; P09616; 7AHL.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
DR Hemolysis; Toxin; Signal; Complete proteome.
KW SIGNAL 1 25
FT CHAIN 26 325
FT SEQUENCE 325 AA; 36711 MW; 082999EB9339B6A0 CRC64;
SQ
Query Match 27.2%; Score 404; DB 1; Length 325;
Best Local Similarity 29.4%; Pred. No. 1e-18;
Matches 87; Conservative 74; Mismatches 119; Indels 16; Gaps 6;
QY 9 TGFTDI-GSNTTV---KTGDLVITYDKENG-----MHKKVYFSDIDKNNHKKLLVIR 56
DB 21 SGTANAEGKITPPVSVKVDKVTLYKTATADSKFKISQILTFNFKDKSYDKDTLVLK 80
QY 57 TGTIAQGVRYVEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDFKYXSL 116
DB 81 ATGNISGFVKPNPDYDFSKLWGAKNVSISSQNSDVSNNVVDYAPKQNEEFQVNTL 140
QY 117 XYFXGXVGDGXGXIXGXIXAXVIXHXLXYXQDPFKTILESPD-KKVGWKFVFNMV 175
DB 141 GYTFGGDI--SISNGLSGGLNGTAFSETINYKQESYRTLSRNTNKNVWGVEAHKIM 198
QY 176 QNWGPDYDRSWPNVGNQFMKTRNGSKMAADNFDLPNKASSLLSGSPDPFATVITMD 235
DB 199 NNGWPGYGRDSFHPYTGNEFLAGROSSAYAGQNFIAQHQMPLLSRSNFEPLSVLS-H 257
QY 236 RKASKQOTNDIVYERVDYQLHWTSTNKGNTKDKWTDSSRSYKIDWEKEEM 291
DB 258 RQDGAKKSKITVYQREMDLYQIRWNGFYWAGANYKNFKTRFKSYEIDWENHKV 313
RESULT 3
LUKF_STAAU STANDARD; PRT; 323 AA.
AC P31715;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leukocidin F subunit precursor (Gamma-hemolysin, H-gamma-I subunit).
GN LUKF
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;

```

```

RN SEQUENCE FROM N.A., AND SEQUENCE OF 26-66.
RP STRAIN=MRSA NO. 4;
RX MEDLINE=92246945; PubMed=1575738;
RA Rahman A., Nariya H., Izaki K., Kato I., Kamio Y.;
RT "Molecular cloning and nucleotide sequence of leukocidin F-component
RT gene (lukF) from methicillin resistant Staphylococcus aureus.";
RL Biochem. Biophys. Res. Commun. 184:640-646(1992).
RN [2]
RP SEQUENCE OF 27-85 AND 322-323.
RX STRAIN=RIMD 310925;
RC MEDLINE=93223836; PubMed=8467905;
RA Kamio Y., Rahman A., Nariya H., Ozawa T., Izaki K.;
RT "The two Staphylococcal bi-component toxins, leukocidin and gamma-
RT hemolysin, share one component in common.";
RL FEBS Lett. 321:15-18(1993).
CC -!- FUNCTION: LEUKOCIDIN CAUSES CYTOTOXIC CHANGES IN POLYMORPHONUCLEAR
CC LEUKOCYTES. GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.
CC -!- SUBUNIT: LEUKOCIDIN CONSISTS OF TWO PROTEIN COMPONENTS: F AND S;
CC GAMMA-HEMOLYSIN CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I-F
CC AND H-GAMMA-II).
CC -!- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S65052; AAC60446.1; -.
DR PIR; J01530; JQ1530.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
KW Hemolysis; Toxin; Signal.
FT SIGNAL 1 25
FT CHAIN 26 323
FT SEQUENCE 323 AA; 36567 MW; A5951CF2AAB6C7E8 CRC64;
SQ
Query Match 26.3%; Score 390; DB 1; Length 323;
Best Local Similarity 28.8%; Pred. No. 7.9e-18;
Matches 85; Conservative 74; Mismatches 120; Indels 16; Gaps 6;
QY 9 TGFTDI-GSNTTV---KTGDLVITYDKENG-----MHKKVYFSDIDKNNHKKLLVIR 56
DB 21 SGTANAEGKITPPVSVKVDKVTLYKTATADSKFKISQILTFNFKDKSYDKDTLVLK 80
QY 57 TGTIAQGVRYVEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDFKYXSL 116
DB 81 ATGNISGFVKPNPDYDFSKLWGAKNVSISSQNSDVSNNVVDYAPKQNEEFQVNTL 140
QY 117 XYFXGXVGDGXGXIXGXIXAXVIXHXLXYXQDPFKTILESPD-KKVGWKFVFNMV 176
DB 141 GYTFGGDI--SISNGLSGGLNGTAFSETINYKQESYRTLSRNTNKNVWGVEAHKIM 198
QY 177 QNWGPDYDRSWPNVGNQFMKTRNGSKMAADNFDLPNKASSLLSGSPDPFATVITMDR 236
DB 199 -GWGPGYGRDSFHPYTGNEFLAGROSSAYAGQNFIAQHQMPLLSRSNFEPLSVLS-HR 256
QY 237 KASKQOTNDIVYERVDYQLHWTSTNKGNTKDKWTDSSRSYKIDWEKEEM 291
DB 257 QDRAKSKITVYQREMDLYQIRWNGFYWAGANYKNFKTRFKSYEIDWENHKV 311
RESULT 4
HLGA_STAAU STANDARD; PRT; 309 AA.
ID P31714; Q07225; Q53689; Q53690;
AC

```



```

Db 33 TEDIGKSDI---EIKRTEDKTSNKGWYQNIQFVADTKYNDALILKMQGFSSRT 89
QY 66 RYSEEGANK-SGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKXXYXSLXYFXGXV 124
Db 90 TYINXKTHVAMRWFOYINIGLNTDNDKYVSLI-NYLPKNKIESTNSQTLGYNIGNF 148
QY 125 XGDXGXIXGXIXAXVIXHXLXYYQDPFKTITLESPTDKKVGKVFNNQNGPYDR 184
Db 149 QASPSLGGNGSF---NYSKSIYQOQNVSEVEQNSKSVLWGYKANSFATES----- 198
QY 185 DSWNPYGNQLPMKTRNGSKAADNPLDPNKASSLLSSGFSDFATVITMDRKASKQQTN 244
Db 199 -GQSAFSDLEFGVGYKPHSKDRDYFVDPSELPLVQSGFNPSFIATVSHE-KGSSDTSE 256
QY 245 IDVIVERVDR---DYQLHWTSTNKGKNTKDKWTDTRS-SERYKIDMEKEEM 291
Db 257 FEITGRNMDVTHAKRSTHYGNSYLDGHRVHNAFVNRNITVYKVEVNNWKTHEI 309

```

RESULT 7

```

ID YM81 YEAST STANDARD; PRT; 1562 AA.
AC Q04781; Q04029;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 180.2 kDa protein in FAA4-HOR7 intergenic region.
GN YMR247C OR YH9408.09C OR YH9920.01C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 1-956 FROM N.A.
RC STRAIN=S288c / AB972;
RA Hunt S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 950-1562 FROM N.A.
RC STRAIN=S288c / AB972;
RA Genties S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -!- SIMILARITY: STRONG, TO HUMAN ZNF294.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 248639; CAA88574.1; -
DR EMBL; 248756; CAA88657.1; -
DR SGD; S0004861; YMR247C.
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; 2F_RING_1; FALSE_NEG.
DR PROSITE; PS00089; 2F_RING_2; 1.
KW Hypothetical protein: Zinc-finger.
FT ZN_FING 1508 1555 RING-TYPE.
SQ SEQUENCE 1562 AA; 180185 MW; 97AC65E881362305 CRC64;

```

```

Query Match      8.1%; Score 121; DB 1; Length 1562;
Best Local Similarity 19.3%; Pred. No. 21;
Matches 62; Conservative 74; Mismatches 106; Indels 80; Gaps 16;
QY 30 KENGHHKVFVFDIDK-----NHNKKL-----LVTRTGTIAGQY--RYVSEEGANK 75
Db 1166 KLESQYKRIFEVLNDDKDGNSRLLLTILGSLVYVTKQDIIIEYELRIQKGTGSDV 1225

```

```

QY 76 SGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKXXYXSLXYFXGXV----- 124
Db 1226 DGSASDNDVNSFKLPQLKLOKVTDEVKPEYLEYENKNSFIKYLAWHLLLMYFKDTSYN 1285
QY 125 -----XGDXGXI---XGIXAXVIXHXLXYYQDPFKTITLESPTDKKVGKVFNNM 174
Db 1286 MRQIFIEQLKEAGLINRMFDITDQIDLRDTEFWQVDTKEISE-----YNI 1332
QY 175 VNQNGPGYDRDSW---NPVYGNQLEPMKTRNGS-----MKAADNPLDPNKASSLLSSGF 224
Db 1333 VGNFNSPKYEDIFECKLLGLTLYQLFNNVNGCLTSIWNLIKDKRTLQ-NDIEKEVVSFEI 1391
QY 225 SP-----DFATVIT-MDRKASKQQT---NIDVIYERVR-----DDYQLHWT---STNWK 267
Db 1392 SPILIKNEFDDINSKMDRLTSNDALTIKLNITNEVRASYLIDDOKLEISFKLPKNYPL 1451
QY 268 TNTKDKWTDNR---SSERYKIDW 286
Db 1452 TNIOVNGVSRVIGSEQKWK-QW 1472

```

RESULT 8

```

ID CHD3_CAEEL STANDARD; PRT; 1787 AA.
AC Q22516; Q18794;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chromodomain helicase-DNA-binding protein 3 homolog (CHD-3).
GN CHD-3 OR T1468.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Matthews P., Murray A.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Chromatin-remodelling protein that function in vulval
CC cell fate determination.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
CC -!- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
CC -!- SIMILARITY: CONTAINS 2 CHROMO DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF308444; AAC29837.1; -
DR EMBL; Z67884; CAA91810.1; -
DR EMBL; Z67881; CAA91810.1; JOINED.
DR EMBL; Z67881; CAA91798.1; -
DR EMBL; Z67884; CAA91798.1; JOINED.
DR WormPep; T1468.1; CE03657.
DR InterPro; IPR000953; Chromo.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002464; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000330; SNF2_N.

```

```

DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00176; SNF2_N.1.
DR Pfam; PF00271; helicase_C.1.
DR Pfam; PF00385; chromo.1.
DR Pfam; PF00628; PHD.2.
DR SMART; SM00298; CHROMO.2.
DR SMART; SM00487; DEXDC.1.
DR SMART; SM00490; HELICC.1.
DR SMART; SM00249; PHD.2.
DR SMART; SM00184; RING.2.
DR PROSITE; PS00598; CHROMO.1; FALSE_NEG.
DR PROSITE; PS00113; CHROMO.2.1.
DR PROSITE; PS00690; DEAH_ATP_HELICASE.1.
DR PROSITE; PS01359; ZF_PHD.1.2.
DR PROSITE; PS00116; ZF_PHD.2.2.
DR Chromatin regulator; Nuclear protein; Repeat; Helicase; DNA-binding;
KW ATP-binding; Zinc-finger.
FT DOMAIN 59 62 POLY-LYS.
FT ZN_FING 265 312 PHD-TYPE 1.
FT ZN_FING 328 375 PHD-TYPE 2.
FT DOMAIN 373 476 CHROMO 1.
FT DOMAIN 501 583 CHROMO 2.
FT DOMAIN 1287 1291 POLY-ARG.
FT NP_BIND 641 648 ATP (POTENTIAL).
FT SITE 763 766 DEAH BOX.
SQ SEQUENCE 1787 AA; 205254 MW; 1EFCE1FECE59740 CRC64;

Query Match 7.6%; Score 113.5; DB 1; Length 1787;
Best Local Similarity 17.6%; Pred. No. 79;
Matches 63; Conservative 69; Mismatches 129; Indels 97; Gaps 15;

QY 7 IGTGTTDIGNTVTKGDLVYDK-----ENGMHKKVFS-----FIDDKHNK 50
DB 925 LKNGMYE-GSALKKNAGKVFLLQKMLKLDGHRVLFISQMTMLDILEDKDFCVGYKY 93
QY 51 KLIVRTKGTIAGYRV-----YSEEGANK-----SGLAWPSAFKVLQLPNEVA 96
DB 984 E-----RIDGSITGQRODAIDRYNAPGAKQFVLLSTRAGGLG-----INLATADTVII 1033
QY 97 QISDYPRNSIDT-----KKYKSLXXF-----XGVVGCXD 129
DB 1034 YSDWPHNDIQAFSRAHRLGQKHVMYIRFVTKGSVEERITSVAKKMLLTHLVVRAGL 1093
QY 130 GXIXGXIXAXVHXLYXQPD-FK-----TILESPDKKVGWK-----V 169
DB 1094 GAKDGKSMSTELDDVLRWGTEELFKEEAPVEGADGEGTSKKPNEQEIWDDAAVDL 1153
QY 170 IFNNMVNQNGPYDRDSWNPVYGNQLFMKTRNGSMKAADNFDLPNKASLLSGF-----SP 226
DB 1154 LDRNKEEGQDGKKEHTWTEYLSFSKVATYN--TKAADADDEDETEVIKGTEDQDP 1211
QY 227 DFATVTIMDRKASKQQTNIIDVIRVDRDYQLHWTSN-----WKGNTDKWTDGRSE 280
DB 1212 NYWEKLLKHYYEQDETQELQKLGKRRVQVYASENMGQWSAQNNOOEEEDGSE 1269

RESULT 9
Y943_METJA
ID Y943_METJA STANDARD; PRT; 417 AA.
AC Q83353;
DT 15-JUL-1998 (Rel. 36, Created)
DE 15-JUL-1998 (Rel. 36, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0943.
GN MJ0943.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;

```

```

RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.F., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.G., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Kleink H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67538; AAB98953.1; -.
CC TIGR; MJ0943; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 148 168 POTENTIAL.
FT DOMAIN 399 415 PRO-RICH.
FT SEQUENCE 417 AA; 46355 MW; 6DF8D82E1AA8BD90 CRC64;

Query Match 7.6%; Score 113; DB 1; Length 417;
Best Local Similarity 19.9%; Pred. No. 87;
Matches 56; Conservative 64; Mismatches 113; Indels 48; Gaps 12;

QY 4 DINIKGTDTGNTTVTKGDLVYDKENGMHKKVFSFIDDKHNHKKLLVIRTKGTIAG 63
DB 34 DININTNN-GENTE-KPINQENNVNENKESQSTQNIQSYENKEIKNQENHPLOS 91
QY 64 QYRV-----YSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYKSLX 117
DB 92 NQNYEQTNNGNFENENAMTNVGESEVNNPEAYNYIET--YPDGTIPDKIEQMLY 149
QY 118 Y-----XFXGXVXGD---XGXIXGXIXAXVHXLYXQPDKFKTILESPTDKKVGWK 169
DB 150 YIKVIDPIVGLAGIDYVDGNYIGTLDVYGVGECVFY-EPGYHTITAEADNGKILASKT 208
QY 170 I-----FNNMVNQNGPYDRDSWNPVYGNQLFMKTRNGSMKAADNFDLPNKASLL-- 220
DB 209 VYVEEGTAYNSGESENDEYDNN-----YESNDLQQTQT-QFSEIEVYVDIIRPSNIII 262
QY 221 -----SSGFSPPDFATVITMD--RKASKQQTNIIDV 247
DB 263 TKLAMPNGFLASINGISPDIGVNIEMENGEKINLKYVMDV 303

RESULT 10
AMY_BUTFI
ID AMY_BUTFI STANDARD; PRT; 976 AA.
AC P30269;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN AMYA.
OS Butyrivibrio fibrisolvens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
OC Butyrivibrio.
OX NCBI_TaxID=831;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HI7C;
RX MEDLINE=91286207; PubMed=2061294;

```

RA Rumbak E., Rawlings D.E., Lindsey G.G., Woods D.R.;
 RT "Cloning, nucleotide sequence, and enzymatic characterization of an
 RT alpha-amylase from the ruminal bacterium Butyrivibrio fibrisolvens
 RT H17c.";
 RL J. Bacteriol. 173:4203-4211(1991).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M62507; AAA23005.1; -
 DR PIR: A42466; A42466.
 DR HSP: P00691; IBAG.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR005085; CBM_25.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase; 1.
 DR Pfam: PF03423; CBM_25; 1.
 KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Signal.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 976 ALPHA-AMYLASE.
 FT ACT_SITE 323 323 BY SIMILARITY.
 FT ACT_SITE 327 327 BY SIMILARITY.
 FT ACT_SITE 447 447 BY SIMILARITY.
 SQ SEQUENCE 976 AA; 106695 MW; 61D690FDC19315AC CRC64;
 Query Match 7.5%; Score 112; DB 1; Length 976;
 Best Local Similarity 20.7%; Pred. No. 38;
 Matches 56; Conservative 47; Mismatches 75; Indels 92; Gaps 15;
 QY 10 GTTDTGNTVTKTGLVYDENGKMHKVFYFDKNNKLLVIRKGTAGQYRVYS 69
 DB 276 GYTDRLTYISMGPLDVPDTEGTFQYFVFLKD-----CVYL 315
 QY 70 EGAANKSGLAWPSAFK---QLQLPDNEVAQISDY-----YPR-----NSID 108
 DB 316 --GAD-----GPRDTAKHISLPDPPV--SDYDAGRNTFYPNNREALNEYSEVG 363
 QY 109 TKYXASXLXY-----XFXGVXGXDXG-----XIXGXIXAXVIXHXIX 147
 DB 364 TKSVDLFPYGVGLGTNDRLAAVQYIGGTTASNYGSSLSRSLSGNLSVNRLLDYQI- 422
 QY 148 YQPDFFKILSPDTPDKVGVKVIENNQN---WGPYDRD-----WNPV-----YGNOLF 196
 DB 423 YDDTAYGSTYATDEKLTWYVESHNDYMNDESCKWISDDDMVINGWSIIAARDAGTPLF 482
 QY 197 MKTRNGSKMAADFDPNKAASSLLSSGFSP 226
 DB 483 FSRPNNS--SAENPYGDN----LIGAGSP 506
 RESULT 11
 ID_CYSF_PLAFA STANDARD; PRT; 569 AA.
 AC P25805;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trophozoite cysteine proteinase precursor (EC 3.4.22.-) (TCP).
 OS Plasmodium falciparum
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OK NCBI_TaxID=5833;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=9228005; PubMed=1565129;
 RX

RA Rosenthal P.J., Nelson R.G.;
 RT "Isolation and characterization of a cysteine proteinase gene of
 RT Plasmodium falciparum.";
 RL Mol. Biochem. Parasitol. 51:143-152(1992).
 CC -1- FUNCTION: PROBABLY DEGRADATES ERYTHROCYTE HEMOGLOBIN.
 CC -1- DEVELOPMENTAL STAGE: TROPHOZOITE
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M81341; AAA29578.1; -
 DR PIR: A45624; A45624.
 DR HSP: P07686; 10DQ.
 DR MEROPS: C01.077; -
 DR InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR000169; SHprot_acsite.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR00705; PAPAIN.
 DR ProDom: PD000158; Peptidase_C1; 1.
 DR PROSITE: PS00139; THIOLEPROTEASE_CYS; 1.
 DR PROSITE: PS00639; THIOLEPROTEASE_HIS; 1.
 DR PROSITE: PS00640; THIOLEPROTEASE_ASN; 1.
 KW Hydrolase; Thiol protease; Zymogen; Glycoprotein; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT PROPEP ? 332 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 333 569 TROPHOZOITE CYSTEINE PROTEINASE.
 FT DOMAIN 64 70 POLY-ASN.
 FT ACT_SITE 357 357 BY SIMILARITY.
 FT ACT_SITE 488 488 BY SIMILARITY.
 FT ACT_SITE 533 533 BY SIMILARITY.
 FT DISULFID 354 395 BY SIMILARITY.
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 479 479 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 487 487 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 569 AA; 66880 MW; E5E762A3637F7C60 CRC64;
 Query Match 7.5%; Score 111.5; DB 1; Length 569;
 Best Local Similarity 17.7%; Pred. No. 18;
 Matches 69; Conservative 64; Mismatches 99; Indels 157; Gaps 17;
 QY 10 GTTDTGNTVTKTGLVYDENGKMHKVFYFDKNNKLLVIRKGTAGQYRVYS 69
 DB 198 GALEIGNMEEMK-----YKEDPNNIKYASKFFKFMKEHNKVKYKNIDEQMRKFEIFKI 251
 QY 41 SFIDDKNNK--KLLVIRKGTAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQ- 97
 DB 252 NYSIKNNKLNKNMYKKK---VNOFSDYSEELKE-----YFKTLHVPNHHMEKY 301
 QY 98 -----ISDYPRNSIDTKXYXS-----XLYFXGVXGXDXGXIXGXIXAXV 140
 DB 302 SKPFENHLKDNILISEFTYNGKRNEKIDFSKVPEILDYREKGIHVHEPKDQGLGCSWAF 361
 QY 141 XIXHXLYXQPDFKTIESTPTDKKVGKVIENNQN---WGPYDRD-----WNPV-----YGNOLF 196
 DB 362 SVGNIESVFAKKNKNILSFSEQEVV-----DCSKDNFG---CDGHPFYSFLYLQNE 411
 QY 192 ---GNOLFMTNRN-----GSMKAADNFDLPNKAASSL-LSSGFSDFDA 229
 DB 412 LCIGDEYKAKDDMFCLNRCRKRKVSLSIGAVKENQLILALNEVGPLSVNNGVNDNFV 471
 QY 230 T-----VITMDRKASKQQTNDIVYERVR 253
 DB 472 AYSEGVYNGTCSELNHSLVLLVGYGVEKTKLNNKNIQTYNTKENSQPDNIIY---- 527

```
QY 254 DDYQLHWTSTN-WKGTNTKDKWTRDSSER 281
Db 528 -----YWIKNWS-----SKKKGNGFMR 546

RESULT 12
HX2_HAEIN STANDARD; PRT; 928 AA.
AC P4354;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heme/hemopexin-binding protein precursor (Heme:hemopexin utilization
DE protein A).
GN HXUA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-35.
RC STRAIN=DL42 / Serotype B;
RX MEDLINE=95115556; PubMed=7815944;
RA Cope L.D., Thomas S.E., Latimer J.L., Slaughter C.A.,
RA Mueller-Eberhard U., Hansen E.J.;
RT "The 100 kDa haem:haemopexin-binding protein of Haemophilus
RT influenzae: structure and localization.";
RL Mol. Microbiol. 13:863-873(1994).
RN [2]
RP SEQUENCE OF 1-30 FROM N.A.
RC STRAIN=DL42 / Serotype B;
RX MEDLINE=95270579; PubMed=7751272;
RA Cope L.D., Yorgev R., Mueller-Eberhard U., Hansen E.J.;
RT "A gene cluster involved in the utilization of both free heme and
RT heme:hemopexin by Haemophilus influenzae type b.";
RL J. Bacteriol. 177:2644-2653(1995).
CC -|- FUNCTION: BINDS HEME/HEMOPEXIN COMPLEXES.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U08348; AAA74138.1; -.
KW Transport; Signal; Repeat.
FT SIGNAL 1 21
FT CHAIN 22 928 HEME/HEMOPEXIN-BINDING PROTEIN.
FT DOMAIN 101 679 6 X 6 AA APPROXIMATE REPEATS.
FT REPEAT 101 106 1-1.
FT REPEAT 205 210 1-2.
FT REPEAT 279 284 1-3.
FT REPEAT 410 415 1-4.
FT REPEAT 635 640 1-5.
FT REPEAT 674 679 1-6.
FT DOMAIN 149 172 4 X 6 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 149 154 2-1.
FT REPEAT 155 160 2-2.
FT REPEAT 161 166 2-3.
FT REPEAT 167 172 2-4.
SQ SEQUENCE 928 AA; 101228 MW; 67D45466A4B92390 CRC64;

Query Match 7.5%; Score 111.5; DB 1; Length 928;
Best Local Similarity 19.7%; Pred. No. 38;
Matches 65; Conservative 70; Mismatches 104; Indels 91; Gaps 15;

QY 1 ADSINIKTKTDTIGSNTVTKTGDLYTDKENG-----MHKKVYFSFDDK 46
Db 379 ADIEDNTGTGTGTGTSSFTGTGSLNANGKVNKGNVNISGRH-----IDSF 431
```

```
QY 47 NHNKKLLIVIRTKTIAGQVRYVEEGANKSGLAWPSA-----FKVQLQLPDNEVAQISD 100
Db 432 RGSDDLKLTNKCHI---DINNADIHSKGRLEFFITSLQNEEDFKSNITITDSKI----- 482
QY 101 YYPN-----SIDTKKXYS-----XLXYFXGVXGXDXGXIXGXIXA-----X 139
Db 483 ----NLNGANGMGLGRSVDEKDYDNRWQKTEGSKRKFDFKMSNVEFNQVDDVILAGGFEK 538
QY 140 VIXHXLYXQDPDKT-----ILSPDTDKYGVKVFNNVNONWGPY--- 182
Db 539 VNLDKIVATQTFNYIDGGVSRNGRKYEYGVLD--LDKRTQLSEL--NQORRWGYYIDL 594
QY 183 DRDSWNP-VYGNOLFMTKTRNGSKAADNFDLPNKASLLSSGFSFPDFATVITMDRKASKQ 241
Db 595 ELDNMRAYLYRFDLFATKNTGRSTIKDTEINISNINLNGKGFVHLLAEIKLD----- 648
QY 242 QTNIDVIYERVDYQLHWTSTNWKGTNTK 271
Db 649 NSKIDITFD--KNSQDTLAQTNRLLGMNGK 676

RESULT 13
EPIP_STAEIP STANDARD; PRT; 461 AA.
ID EPIP_STAEIP AC P30199;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermin leader peptide processing serine protease EPIP precursor
DE (EC 3.4.21.-).
GN EPIP.
OS Staphylococcus epidermidis.
OG Plasmid pTu 32.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TU 3298 / DSM 3095;
RX MEDLINE=92155237; PubMed=1740156;
RA Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
RA Goetz F., Entian K.-D.;
RT "Analysis of genes involved in the biosynthesis of lantibiotic
RT epidermin.";
RL Eur. J. Biochem. 204:57-68(1992).
CC -|- FUNCTION: PROTEASE WHICH CLEAVES THE MATURED LANTIBIOTIC FROM THE
CC MODIFIED PREPEPTIDE (PROBABLE).
CC -|- PATHWAY: Epidermin biosynthesis; last step.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X62386; CAA44257.1; -.
DR PIR; S23420; S23420.
DR HSSP; P00782; 2SBT.
DR MEROPS; S08.060; -.
DR InterPro; IPR000209; Peptidase_s8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Signal; Zymogen; Plasmid.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 ? POTENTIAL.
FT CHAIN ? 461 EPIDERMIN LEADER PEPTIDE PROCESSING
FT FT SERINE PROTEASE EPIP.
FT ACT_SITE 149 149 CHARGE RELAY SYSTEM (BY SIMILARITY).
```

```
FT ACT_SITE 194 194 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 402 402 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 461 AA; 51814 MW; 72770F4F8436D906 CRC64;

Query Match 7.5%; Score 111; DB 1; Length 461;
Best Local Similarity 20.9%; Pred. NO. 14;
Matches 60; Conservative 61; Mismatches 120; Indels 46; Gaps 12;

QY 4 DINIKTGTIDGS----NTVTGTDLVYDKENGMMHKVFFSFIDDKNNHKKLLV--IR 56
DB 149 DSGVNSHTDLKINKIVNEVPKNGFRGSENDESNK-----NFEEDKLNHGLTLVAGQIG 203
QY 57 TRGTIAG-----QYRVYSEGANKGLAWPSAFKVLQQLPDNEVAQIS--DYPRNSIDT 109
DB 204 ANGNLKGVNPGVMNVYFSGSKSEMLVWSKGIIDAANDNDVNLVSLGNLIKDQNK 263
QY 110 KYYXSLXYFXGVXGDXGXIXXAXVIXHXL---XYXQPDFKTLSPDTPDKYVG 166
DB 264 KKLRRDEKVDYDALQKAINVAQKGSIVVAAGVNDGINVKVKEINKRNLNSKTSKKV- 322
QY 167 WKVIENNMYNQ--WGPYDRDSNPV--YGNQLF-MKTRNGSKMAADNFDLPNKASS 218
DB 323 ----TDSFANLNNVTVGSIDNDIYSEFSNFGNFIIDLTMTIGGSYK-----LLDKYGGKDA 374
QY 219 LLSGFSFPDFATVITMDRKASKQQTNDIYVERVR-----DDYQL 258
DB 375 WLEKGYMQKQSVLSTSSNGRYIQSGTSLAAPKVSALALEIDKYQL 421

RESULT 14
ID YCF1_ARATH STANDARD; PRT; 1786 AA.
AC P56785;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Hypothetical 213.7 kDa protein ycf1.
OS Arabidopsis thaliana (Mouse-ear cress).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucrodis II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20039611; PubMed=10574454;
RA Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.;
RT "Complete structure of the chloroplast genome of Arabidopsis
thaliana.";
RL DNA Res. 6:283-290(1999).
CC -!- FUNCTION: NOT YET KNOWN.
CC -!- SIMILARITY: BELONGS TO THE YCF1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP000423; BAA84445.1; -.
CC DR EMBL; AP000423; BAA84433.1; -.
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 1786 AA; 213727 MW; CFFD2A4D776D7E5D CRC64;

Query Match 7.3%; Score 108.5; DB 1; Length 1786;
Best Local Similarity 20.4%; Pred. NO. 1.6e+02;
Matches 68; Conservative 61; Mismatches 110; Indels 95; Gaps 16;

QY 5 INIKTGTIDGSNTVTGTDLVY-DKENGMMHK---VFYSFIDDKNNHKKLLVIRTKGT 60
```

```
DB 1092 INIPRITQTLPSTKILDKYIYKNEENGKKNTLFISTIKNLISNKKKM-----1144
QY 61 TAGQYRVYSEGANKGLAWPSAFKVLQQLPDNEVAQISDYPRNSIDTKY---XSIXL 117
DB 1145 ---SYDLCS-----LSQAYVFYKLSQIKVSNFCKLKAVLE 1176
QY 118 YFXGXVVGXDGXIXGXIXAXVIXHXLXXYQPDFKTLSPDTPDKYVGW-----167
DB 1177 YNIC-----ITSFFYKKNIKVFQ-EGHIFHYELNKTFLNSEVQNWKNLRSQYQYNLP 1230
QY 168 KVIENNMYNQNW-GPYDRDS--WNP-----VYGNQLFEMKTRNGSKMAADNFDLPNK---215
DB 1231 QISWRLVTONWKNKINDSLVNLPSLTKEQSYKKKEDYKKQKFFEADSLNPKHNK 1290
QY 216 -----ASSLSSGFSFPDFATVITMDR---KASKQQTNDIYVERVDDYQLHWTST 263
DB 1291 KDSIYNLFYKSIHSTKFNDFMSIGIALDNLVSLFLEKYNIRGMGE-IRHRKYLDWRIL 1349
QY 264 N-----WKGNTNKKDKWTDSSERY-KID 285
DB 1350 NFWFTKVTIETFPWVDTKSKKRYINTKVQYQKID 1383

RESULT 15
ID FAB2_MYCLE STANDARD; PRT; 420 AA.
AC O69473;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 3-oxoacyl-[acyl-carrier-protein] synthase 2 (EC 2.3.1.41) (Beta-
DE ketoacyl-ACP synthase 2) (KAS 2).
GN KASB OR MLI1656 OR MLCBI243.19C.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteriales; Mycobacteriaceae; Mycobacterium.
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -!- FUNCTION: CATALYZES THE CONDENSATION REACTION OF FATTY ACID
CC SYNTHESIS BY THE ADDITION TO AN ACYL ACCEPTOR OF TWO CARBONS
CC FROM MALONYL-ACP (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
CC [acyl-carrier protein].
CC -!- PATHWAY: Fatty acid biosynthesis. Involved in meromycolate
CC extension.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: BELONGS TO THE BETA-KETOACYL-ACP SYNTHASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL023635; CAA19200.1; ALT_INIT.
```


Search completed: March 4, 2003, 10:43:05
Job time : 21 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:40:04 ; Search time 28.5 Seconds
(without alignments)
2118.309 Million cell updates/sec

Title: CLAIM6

Perfect score: 1485

Sequence: 1 ADSDINKTGTDTGSNTTV.....WDRSSRYKIDWEKEETN 293

Scoring table: BLOSUM62Dx

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriapi.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1481	99.7	319	16 Q99UU6	Q99uu6 staphylococ
2	459.5	30.9	412	2 Q05387	Q05387 bacillus ce
3	409	27.5	325	2 Q57227	Q57227 staphylococ
4	404	27.2	327	2 Q54082	Q54082 staphylococ
5	403.5	27.2	327	2 Q93UU8	Q93uu8 staphylococ
6	402.5	27.1	327	16 Q99T54	Q99t54 staphylococ
7	402.5	27.1	332	16 Q931N4	Q931n4 staphylococ
8	401	27.0	325	16 Q931F3	Q931f3 staphylococ
9	400	26.9	325	2 Q50604	Q50604 staphylococ
10	399	26.9	338	16 Q99SN8	Q99sn8 staphylococ
11	396	26.7	325	9 Q80067	Q80067 staphylococ
12	394.5	26.6	326	2 Q54327	Q54327 staphylococ
13	393	26.5	325	2 Q53747	Q53747 staphylococ
14	383	25.8	322	2 Q53704	Q53704 staphylococ
15	382	25.8	322	2 Q9MBN2	Q9mbn2 staphylococ
16	364	24.5	336	2 Q9EVAL	Q9eval bacillus ce

17	333	22.4	265	2 Q937V0	Q937v0 bacillus ce
18	332	22.4	265	2 Q937V1	Q937v1 bacillus ce
19	325.5	21.9	260	2 Q937V2	Q937v2 bacillus ce
20	309	20.8	336	2 Q46308	Q46308 clostridium
21	309	20.8	336	2 Q46181	Q46181 clostridium
22	305	20.5	309	2 Q9L403	Q9l403 clostridium
23	260	17.5	308	2 Q53731	Q53731 staphylococ
24	257.5	17.3	351	16 Q931I5	Q931i5 staphylococ
25	257	17.3	308	2 Q53703	Q53703 staphylococ
26	257	17.3	308	9 Q9MBN3	Q9mbn3 staphylococ
27	255.5	17.2	331	16 Q99SN7	Q99sn7 staphylococ
28	247	16.6	315	2 Q53701	Q53701 staphylococ
29	245	16.5	286	2 Q9AFA9	Q9afa9 staphylococ
30	244	16.4	286	2 Q9AFB0	Q9afb0 staphylococ
31	244	16.4	315	2 Q53691	Q53691 staphylococ
32	244	16.4	315	16 Q99RL1	Q99rl1 staphylococ
33	241.5	16.3	311	2 Q93UU9	Q93uu9 staphylococ
34	241.5	16.3	311	16 Q99T53	Q99t53 staphylococ
35	237	16.0	312	2 Q50603	Q50603 staphylococ
36	237	16.0	312	9 Q80066	Q80066 staphylococ
37	233	15.7	314	2 Q54081	Q54081 staphylococ
38	231	15.6	312	2 Q53746	Q53746 staphylococ
39	201.5	13.6	310	2 Q54326	Q54326 staphylococ
40	147	9.9	28	2 P971A4	P971a4 staphylococ
41	124.5	8.4	217	2 Q93CF5	Q93cf5 lactococcus
42	123.5	8.3	454	16 Q8RGZ8	Q8rgz8 fusobacteri
43	122	8.2	364	2 Q68629	Q68629 haemophilus
44	121	8.1	360	2 Q68626	Q68626 haemophilus
45	121	8.1	572	3 P87247	P87247 botrytis ci

ALIGNMENTS

RESULT 1

ID	Q99UU6	PRELIMINARY;	PRT;	319 AA.
AC	Q99UU6;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Alpha-hemolysin precursor.			
GN	SAV1163 OR SAI007.			
OS	Staphylococcus aureus (strain Mu50 / ATCC 700699), and			
OS	Staphylococcus aureus (strain N315).			
OC	Bacteria; Firmicutes; Bacillus/clostridium group; Bacillales;			
OC	Staphylococcus.			
OX	NCBI_TaxID=158878, 158879;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);			
RX	MEDLINE=21311952; PubMed=11418146;			
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,			
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,			
RA	Kanamori M., Matsumaru H., Murayama A., Murakami H., Hosoyama A.,			
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,			
RA	Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,			
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,			
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,			
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus			
RT	aureus."			
DR	Lancet 357:1225-1240(2001).			
DR	EMBL; AP003361; BAB57325.1; "			
DR	EMBL; AP003132; BAB42258.1; "			
DR	HSSP; P09616; TAAHL.			
DR	InterPro; IPR001776; Aerolysin.			
DR	InterPro; IPR001340; Hemlysn_pore.			
DR	InterPro; IPR003963; Staph_bicn_txn.			
DR	Pfam; PF01117; Aerolysin; 1.			
DR	PRINTS; PR01468; BICOMPENTOXIN.			
DR	TIGRfams; TIGR01002; hlyII; 1.			
DR	PROSITE; PS00274; AEROLYSIN; 1.			
KW	Complete proteome.			

KN	1b]	
KN	SEQUENCE FROM N.A.	
RP	STRAIN=P83;	
RC	MEDLINE=95210788; PubMed=7696748;	
RX	Prevost G., Supersac G., Piemont Y., Colin D., Sire S., Henslen T.,	
RA	Pettau P., Meunier O., Gibier B., Koenig W., Piermont Y.:	
RA	"A particular class of virulence factors: calmodulin-activated	
RT	bacterial adenylate cyclases.";	
RL	Zentralbl. Bacteriol. 24:284-293(1994).	

```
DR EMBL; X81586; CAA57278.1; -.
DR EMBL; X64389; CAA45729.1; -.
DR HSP; P09616; 7AHL.
DR InterPro; IP001776; Aerolysin.
DR InterPro; IP001340; Hemlysn_pore.
DR InterPro; IP003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
KW Signal.
FT SIGNAL.
FT CHAIN
SQ SEQUENCE 325 AA; 36812 MW; 38A0066D8DBD8935 CRC64;

Query Match 27.5%; Score 409; DB 2; Length 325;
Best Local Similarity 28.7%; Pred. No. 1.1e-18;
Matches 81; Conservative 73; Mismatches 112; Indels 16; Gaps 4;

QY 11 TTDTGNTTAKTGLVYDKENGHKKVYFSDIDKHNHKKLVIRTKGTIAGQYRVISE 70
DB 47 TTATADSDKFKISQILT-----FNFIDKSKYDKDTLVLAAGNINSYGERPNP 94
QY 71 EGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSLXLYFXGXVXGXDXG 130
DB 95 KQYDFSKIIYWGAKYNYVSISSQSDNSVNVVDYAPKNQNEEFQVQNTLGYTFGGDI--SISN 152
QY 131 XIXGXIXAXVXIXHXLXIXOPDKFTILESPTD-KKVGKVIFFNMVYQNGWPDYDRDSDNP 189
DB 153 GLSGGLNGTAFSETINYKQESYRTVLSRNTNKNVKGWGVGAHKNMNGWPGYGRDSFHP 212
QY 190 VYGNQLFMKTRNGSKMAADNFDPNPKASSLLSGFSPDFATVITMDRKASKQQTNDIVY 249
DB 213 TYGNELFLAGRSSAYAGONFIAHQHMLLSRNFENPELSVLS-HRQDCAKSKITVTY 271
QY 250 ERYVDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEEM 291
DB 272 QREMDLIQIRNWFYAGANYKFKTRTFKSTYEIDWENHKV 313

RESULT 4
O54082 PRELIMINARY; PRT; 327 AA.
AC O54082;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE Leukotoxin, LukD precursor.
GN LUKD.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NEWMAN;
RC MEDLINE=98452939; PubMed=9781679;
RA Gravet A., Colin D.A., Keller D., Giradot R., Montell H., Prevost G.;
RT "Characterization of a novel structural member, LukE-LukD, of the bi-
component staphylococcal leucotoxins family.";
RL FEBS Lett. 436:202-208(1998).
DR EMBL; Y13225; CAA73668.1; -.
DR InterPro; IP001776; Aerolysin.
DR InterPro; IP001340; Hemlysn_pore.
DR InterPro; IP003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
KW Signal.
FT SIGNAL.
FT CHAIN
SQ SEQUENCE 327 AA; 36876 MW; 6CC5677521C3CD24 CRC64;

Query Match 27.2%; Score 404; DB 2; Length 327;
```

```
Best Local Similarity 28.4%; Pred. No. 2.4e-18;
Matches 84; Conservative 74; Mismatches 114; Indels 24; Gaps 7;

QY 9 TGTDTGSGNTT-----VKTGDLVY-----DKNGMHKKVYFSDIDKHNHKKLLVI 55
DB 21 SNTVDAQNITPKREKKVDDKITLYKTATSDNDKLN-IFQILTFNFKDSYDKDTLVL 79
QY 56 RTKGTIAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSLX 115
DB 80 KAAGNINSYKNSNPKDYNSQFYWGKYNVSVSSSDAVNVVDYAPKNQNEEFQVQOT 139
QY 116 LXYFXGXVXGXDXGXIXXIXAXVXIXHXLXIXQPDFKFTILESPTD-KKVGKVIFFNM 174
DB 140 LGYSYGGDI--NISNGLSGGLNGSKSFSETINYKQESYRTTIDRKNHKSIGWGAH 197
QY 175 VNQWGPYDRDSDNPVYGNQLFPMKTRNGSKMAADNFDPNPKASSLLSGFSPDFATVIT- 233
DB 198 MNGWGPYGRDSYDPTYGNEFLGSGKSSNAGNQLPHTQIPLLARGNFNPEFISVL 257
QY 234 --MDRKASKQQTNDIVYERVRDYLHWTSTNWKGTNTKDKWTDSSERYKIDWE 287
DB 258 KLFDTKSK-----IKTYOREMDRYTNQNRSHWGNKNQNTVFTTSTYEYDQW 309

RESULT 5
Q93U08 PRELIMINARY; PRT; 327 AA.
AC Q93U08;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE LukNF precursor.
GN LUKNF.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RA Morinaga N., Noda M.;
RT "Purification and cloning of novel members, LukNS and LukNF, of
staphylococcal leukocidin family.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055623; BAB47175.1; -.
DR InterPro; IP001776; Aerolysin.
DR InterPro; IP001340; Hemlysn_pore.
DR Pfam; PF01117; Aerolysin; 1.
DR TIGRFAMS; TIGR01002; hlyII; 1.
KW Signal.
FT SIGNAL.
FT CHAIN
SQ SEQUENCE 327 AA; 36889 MW; 3807DB6421ACDAAA CRC64;

Query Match 27.2%; Score 403.5; DB 2; Length 327;
Best Local Similarity 28.8%; Pred. No. 2.6e-18;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;

QY 18 TTVKTGDLVYDKENGHKKVYFSDIDKHNHKKLVIRTKGTIAGQYRVYSEEGANKSG 77
DB 43 TLYKTTATSDNDKLN-ISQILTFNFKDSYDKDTLVLAAGNINSYKKNPKDYNYSQ 101
QY 78 LAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSLXLYFXGXVXGXDXGXIXXIX 137
DB 102 FYWGGKYNVSVSSSDAVNVVDYAPKNQNEEFQVQOTLGYSYGGDI--NISNGLSG 159
QY 138 AXVXIXHXLXIXQPDFKFTILESPTD-KKVGKVIFFNMVYQNGWPDYDRDSDNP 196
DB 160 GSKSFSETINYKQESYRTTIDRKNHKSIGWGAHKNMNGWPGYGRDSYDPTYGNE 219
QY 197 MKTRNGSKMAADNFDPNPKASSLLSGFSPDFATVITMDRKASKQQTNDIVYER 256
DB 220 LGGRQSSNAGQNFLPHTHOMPLIARGNFNPEFISVLSHQNDTK-KSKIKVYOREMD 278
QY 257 QLHWTSTNWKGTNTKDKWTDSSERYKIDWE 287
```



```
Db 161 LAGNITKESNYSETISYQQPSYRTLLDQSTSHKGVGWKVEAHLLNNMGHDHTROLTNDSD 220
QY 188 NPVYGNQLEFMRNGSMKAADNFDLPNKASSLLSSGSPDFATVITMDRKASKQQTINIDV 247
Db 221 NRT-KSEISLFRNGNLWAKDNFTPKDKMPVTVSEGFNPEFLAVMSHDKK-DKGKSQFVV 278
QY 248 IYERVDYQLHWTSTN---WKGNTNKDKWTDSSERYKIDWEKEM 291
Db 279 HYKRSMDPEFKIDNRHGFYWSGENHVDKKEKLSALYEVWDKTHDV 326

RESULT 11
OB0067 PRELIMINARY; PRT; 325 AA.
AC OB0067;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
LUKF-PV.
GN LUKF-PV.
OS Staphylococcus aureus bacteriophage PVL, and
OS Staphylococcus aureus temperate phage phisLT.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=71366; 130478;
RN [1]
RP SPECIES=Phase phi PVL;
RC MEDLINE=98067870; PubMed=9404084;
RA Kaneko J., Kimura T., Kawakami Y., Tomita T., Kamio Y.;
RT "Panton-valentine leukocidin genes in a phage-like particle isolated
RT from mitomycin C-treated Staphylococcus aureus V8 (ATCC 49775).";
RL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Staphylococcus aureus temperate phage phisLT;
RA Narita S., Kaneko J., Chiba J., Etienne J., Piemont Y., Kamio Y.;
RT "Phage conversion of Panton-Valentine leukocidin (PVL) in
RT Staphylococcus aureus: molecular analysis of a PVL-converting phage,
RT phisLT.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB009866; BAA31901.1; -
DR EMBL; AB045978; BAB21755.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemolysin_pore.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMs; TIGR01002; hlyII; 1.
SQ SEQUENCE 325 AA; 36962 MW; 1955C10D28DFCC4F CRC64;

Query Match 26.7%; Score 396; DB 9; Length 325;
Best Local Similarity 27.0%; Pred. No. 7.9e-18;
Matches 75; Conservative 77; Mismatches 110; Indels 16; Gaps 4;

QY 11 TTDIGSNTVKTGDLVTDKENGMMKVFYFIDDKNKKLLVIRTKGTAGQYRVYSE 70
Db 45 FTATSDSKLISQILT-----FNFIKDSYDKDITLKAAGNIYSGYTKNP 92
QY 71 EGANKSLWPSAFKVLQLPDNEVAQISDYIPRNSIDTKYXSLXVXGXGXGXDXG 130
Db 93 KDTISSQFYWSKYNISNSDSNVVVDYAPKNQNEEFQVQTVGVSYGDI--NISN 150
QY 131 XIXGXIXAAXVIXHXLXYXQPDFKLTILESPTD-KVGVKVFNNMNVNNGWGYDRDSWNP 189
Db 151 GLSGGGNGSKSFSETINYKQESYRTSLDKRTNFKIGWDEAHKIMNNGWGYGRDSVHS 210
QY 190 VYGNOLFMRKTRNGSMKAADNFDLPNKASSLLSSGSPDFATVITMDRKASKQQTINIDVIY 249
Db 211 TYGNEWFLGSRQSNLNAGQNFLEHMKPVLGRGNFNPFIQVLSGRKQNAAR-KSKITVY 269
QY 250 EYRDDYQLHWTSTNWKNTNKDKWTDSSERYKIDWE 287
Db 270 QREMDRYTFNWNQLHWIGNNYKDENRATHTSIYVDWE 307
```

```
RESULT 12
Q54327 PRELIMINARY; PRT; 326 AA.
AC Q54327;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Synergohymenotropic toxin.
GN LUKF.I.
OS Staphylococcus intermedius.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B62;
RX MEDLINE=96105366; PubMed=7498527;
RA Prevost G., Bouakham T., Piemont Y., Montell H.;
RT "Characterisation of a synergohymenotropic toxin produced by
RT Staphylococcus intermedius.";
RL FEBS Lett. 376:135-140(1995).
DR EMBL; X79188; CAA55783.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemolysin_pore.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMs; TIGR01002; hlyII; 1.
FT CHAIN 27 326 SYNERGOHYMENOTROPIC TOXIN.
SQ SEQUENCE 326 AA; 36597 MW; F8A15CEB345A22B4 CRC64;

Query Match 26.6%; Score 394.5; DB 2; Length 326;
Best Local Similarity 29.6%; Pred. No. 1e-17;
Matches 81; Conservative 76; Mismatches 108; Indels 9; Gaps 8;

QY 18 TTVTGTGDLVTDKENGMMKVFYFIDDKNKKLLVIRTKGTAGQYRVYSEGANKS 76
Db 43 TLYKT--TATADSDNLNISQLLTFFNFIKDSYDKDITLVLKAAGNINSYKSPNPNDIYS 100
QY 77 GLWPSAFKVLQLPDNEVAQISDYIPRNSIDTKYXSLXVXGXGXGXGXIXXI 136
Db 101 SFYWGAKYNSISAESKGAENVVDYAPKNQNEEFQVQNTLGYSGGDIS-ISKG-LSGGL 158
QY 137 XAXVXHXHXLXYXQPDFKLTILESPTDKK-VGVKVFNNMNVNNGWGYDRDSWNPVYNOL 195
Db 159 NGSESFSETINYKQESYRTSLDKRTNFKIGWDEAHKIMNAGWGYGRDSFHDLYGNEL 218
QY 196 FMKTRNGSMKAADNFDLPNKASSLLSSG-FSPDFATVITMDRKASKQQTINIDVIYERVD 254
Db 219 FLGGRQSKLNAGQNFL-PTRQMPLLARGNFNPEFLSVLSHKNPAGAK-TSKIKVTVQREMD 276
QY 255 DYQLHWTSTNWKNTNKDKWTDSSERYKIDWEK 288
Db 277 EYTNWNGFHWGMGTNYKNQNNATFTSFYEIDWDQ 310

RESULT 13
Q53747 PRELIMINARY; PRT; 325 AA.
AC Q53747;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Leucocidin F precursor.
GN LUKF.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE OF 1-285 FROM N.A.
```



```
RC STRAIN-ATCC 49775;
RA Prevost G., Supersac G., Piemont Y., Colin D., Sire S., Henslen T.,
RA Petiau P., Meunier O., Gibier B., Koenig W., Piermont Y.;
RT "The new family of leukotoxins from S.aurea: structural and
RT biological properties.";
RL Zentralbl. Bakteriol. 24:284-293(1994).
[2]
RN SEQUENCE FROM N.A.
RP SEQUENCE ATCC 49775;
RC STRAIN-ATCC 49775;
RX MEDLINE=96009775; PubMed=7558328;
RA Prevost G., Couppie P., Petiau P., Supersac G., Finck-Barbancon V.,
RA Montell H., Piemont Y., Gibier B.;
RT "Panton-Valentine leukocidin and gamma-hemolysin from
RT staphylococcus aureu ATCC 49775 are encoded by distinct loci and have
RT different biological activities.";
RL Infect. Immun. 63:4121-4129(1995).
DR EMBL: X72700; CAA51252.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
KW Signal.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN 25 325 POTENTIAL.
SQ SEQUENCE 325 AA; 37009 MW; 03C2C247B3B3CCD9 CRC64;

Query Match 26.58; Score 393; DB 2; Length 325;
Best Local Similarity 27.08; Pred. No. 1.2e-17;
Matches 75; Conservative 76; Mismatches 111; Indels 16; Gaps 4;

Qy 11 TTDTGSNTTVKGTGLVTVYDKENGHMKVVFSDDKNHNKLLVIRTKTAGOYRVYSE 70
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 TTATSDSKLKSQILT-----FNFIDKSYDKDRTLKKAAGNIYSYTRP 92

Qy 71 EGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSLXLYFXGXVXGDXG 130
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 KDTISSQYMGSKYKYNISNDSNDVNVVDYAPKNQNEEFQVQTVGVSYGGDI--NISN 150

Qy 131 XIXGXIXAXVIXHXLXYPDFKILESPTD-KKVGKVFNNVQNGWPGYDRDSWNP 189
::: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 GLSGGGNGSKSFSETINYKQESYETSLDKRTNFKKIGWDVEAHKIMNMGWPGYGRDSYHS 210

Qy 190 VYGNQLFMKTRNGSKAADNFDLPNKASSLSGSGFPDFATVITMDRKASKQOQTNIDVIY 249
||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 211 TYGNEMFLGSQSNLAGQNFLEHKKMPVLSRGNFNEFEFGLVSRKQNAK-KSKITVY 269

Qy 250 ERVRDDYQLHWTSTNWKGTNTKDKWTRSSRYKIDWE 287
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 270 QSEMDRYTNWFNFWIGNKNYKDHIRATHHTSIYEDWE 307

RESULT 14
Q53704 ID Q53704 PRELIMINARY; PRT; 322 AA.
AC Q53704;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE LukF-PV like component.
GN LukF-PV.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=P83;
RA Kaneko J., Choorit W., Kamio Y.;
RT "LukF-PV like component of leukocidin and gamma-hemolysin in S.aureus
RT P8 3.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D83951; BAA12148.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
SQ SEQUENCE 322 AA; 36519 MW; 9424E58715B108B6 CRC64;

Query Match 25.88; Score 383; DB 2; Length 322;
Best Local Similarity 26.38; Pred. No. 5.6e-17;
Matches 80; Conservative 78; Mismatches 122; Indels 24; Gaps 4;

Qy 6 NIKGTGTDIGSNTTVKGTGLVTVYDKENGHMKV-----FVSFIDD 45
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10 SVATSIILMLSNVDAQAQHTTPVSEKKVDDKITLYKTATSDSKLKSQILTFFIKD 69

Qy 46 KNHNKLLVIRTKTAGOYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYTPRN 105
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 KSYDKDRTLKKAAGNIYSYGTPTSDSINSQFYWGAKYNVFVSESKDSYNIYDAPKN 129

Qy 106 SIDTKYXSLXLYFXGXVXGDXGXIXAXVIXHXLXYPDFKILESPTD-KK 164
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 QNEEFQVQTVGVSYGGDI--NIINGLTGGLNGSKSFSETINYKQESRTTIDRKTNHS 187

Qy 165 VGVKVFNNVQNGWPGYDRDSWNPVYCNQLFMKTRNGSKAADNFDLPNKASSLSGSG 224
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 IGWGEAHKIMNMGWPGYGRDSSDSLYGNELFLGGRSSSNANQNFLEPHOMPILARGNF 247

Qy 225 SPDFATVITMDRKASKQOQTNIDVIYERVDDYQLHWTSTNWKGTNTKDKWTRSSRYKI 284
||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 NPEISVLSHKQDKV-KSKIKVYQREMDRYENFNNLHWIGYNIKNIKRATHTSIYEI 306

Qy 285 DWK 288
||||
Db 307 DWK 310

RESULT 15
Q9MBN2 ID Q9MBN2 PRELIMINARY; PRT; 322 AA.
AC Q9MBN2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE LukF-PV(P83).
OS Staphylococcus aureus prophage phiPV83.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=129009;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P83;
RA Zou D., Kaneko J., Narita S., Kamio Y.;
RT "Complete nucleotide sequence and molecular characterization of
RT prophage PV83pro carrying lukF-pv(p83) gene cluster in
RT Staphylococcus aureus strain P83.";
RN Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=P83;
RX MEDLINE=98067870; PubMed=9404084;
RA Kaneko J., Muramoto K., Kamio Y.;
RT "Gene of LukF-pv-like component of Panton-Valentine leukocidin in
RT Staphylococcus aureus P83 is linked with lukM.";
RL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).
DR EMBL; AB04554; BAA97867.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:38:18 ; Search time 33 seconds
(without alignments)
1183.104 Million cell updates/sec

Title: CLAIM7

Perfect score: 1484

Sequence: 1 ADSIDINIKTGTIGSNTTV.....WDRSSERYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1484	100.0	293	20	AAV01365
2	421.5	28.4	330	22	AAU33797
3	421.5	28.4	333	22	AAU36528
4	395.5	26.7	327	22	AAU34405
5	395.5	26.7	327	22	AAU37136
6	395.5	26.7	327	22	AAU37564
7	388	26.1	325	22	AAU34287
8	388	26.1	325	22	AAU37100
9	388	26.1	325	22	AAU37186
10	388	26.1	325	22	AAU37477

11	388	26.1	325	23	AAU5478	S. aureus antigeni
12	383	25.8	323	22	AAU34395	Staphylococcus aur
13	383	25.8	338	22	AAU37509	Staphylococcus aur
14	374	25.2	323	14	AAU35084	(Pro)leukocidin F.
15	253	17.0	321	22	AAU34275	Staphylococcus aur
16	253	17.0	321	22	AAU37101	Staphylococcus aur
17	253	17.0	321	22	AAU37128	Staphylococcus aur
18	242	16.3	315	14	AAU35083	(Pro)leukocidin S.
19	240	16.2	312	22	AAU33771	Staphylococcus aur
20	133.5	9.0	826	22	AG81986	S. epidermidis epi
21	133.5	9.0	10182	23	ABP38314	Staphylococcus epi
22	132.5	8.9	1300	22	AAU33407	Enterococcus faeca
23	132.5	8.9	1300	22	AAU35313	Enterococcus faeca
24	129	8.7	40	20	AAV01366	Alpha haemolysin (
25	124.5	8.4	1231	20	AAV00219	Enterococcus faeca
26	124.5	8.4	1231	23	ABP43438	E faecalis Eri08 a
27	124.5	8.4	1265	20	AAV00218	Enterococcus faeca
28	124.5	8.4	1265	23	ABP43437	E faecalis Eri08 p
29	118	8.0	447	23	ABO9643	Mutp protein of th
30	117.5	7.9	2478	22	AAU34320	Staphylococcus aur
31	117.5	7.9	2478	22	AAU37374	Staphylococcus aur
32	117	7.9	978	22	AAU33960	Staphylococcus aur
33	117	7.9	1001	22	AAU37093	Staphylococcus aur
34	116.5	7.9	1249	22	ABG21666	Novel human diagno
35	116	7.8	384	15	AAU47236	Wild-type Feline H
36	116	7.8	435	23	ABP55618	Lactococcus lactis
37	116	7.8	2893	19	AAW98828	H. pylori GHPO 148
38	116	7.8	2893	19	AAW71556	Helicobacter poly
39	115.5	7.8	675	22	ABH71143	Drosophila melanog
40	115	7.7	461	14	AAU39354	Epip protein. Sta
41	115	7.7	461	20	AAU43441	S. epidermis readi
42	115	7.7	889	22	AAU45918	S. enterica serova
43	115	7.7	1537	22	ABP58058	Drosophila melanog
44	114.5	7.7	674	22	AAU92775	Human protein sequ
45	114.5	7.7	2368	22	AAU34139	Staphylococcus aur

ALIGNMENTS

RESULT 1

AAV01365

ID AAV01365 standard; protein; 293 AA.

AC AAV01365;

XX

DT 03-JUN-1999 (first entry)

XX

DE Wild-type Staphylococcal alpha haemolysin (HL) polypeptide.

XX

DE Staphylococcal; alpha-haemolysin; alphaHL; mutant; metal; biosensor;
heptameric pore assembly; micronutrient analysis; industrial effluent;
organic compound; explosive; macromolecule; bacteria; virus.

XX

OS Staphylococcus aureus.

XX

PN WO9905167-A1.

XX

PD 04-FEB-1999.

XX

PF 24-JUL-1998; 98WO-US15354.

XX

PR 25-JUL-1997; 97US-0053737.

XX

XX (UYMA-) UNIV MASSACHUSETTS.

XX

PI Bayley H, Braha O, Gouaux E, Kasianowicz J;

XX

DR WPI; 1999-153311/13.

XX

PT New mutant staphylococcal alpha-haemolysin - comprises a
heterologous amino acid that binds to analyte, particularly metal
ions

Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKTGDLVTDKENGHKKVYFSDIDKNNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
| | | | | : : : : : | | | | | : : : : : | | : : : |
DB 43 TLYKTTATSDNDKLN-ISOILTFNFKDKSYDKDTLVLKAAGNINSYGKKPNKPDYNSQ 101
| | | | | : : : : : | | | | | : : : : : | | : : : |
QY 78 LAMPFAFKVQLQDPNEVAQISDYPRNSIDTXEMXNTYTXGXNNXNTYDXTXKXGLXG 137
| | | | | : : : : : | | | | | : : : : : | | : : : |
DB 102 FYWGGKYNVSSESNDVAVNVYAPKN--QNEEFQVQOTLGYSYG--GDINISNGLSGG 157
| | | | | : : : : : | | | | | : : : : : | | : : : |
QY 138 XN--XSXGXTXXVQDPFKTILESTPD-KKYGWKVIFNNVNNQNGPYDRDSWNPVYGNQ 194
| | | | | : : : : : | | | | | : : : : : | | : : : |
DB 158 LNSGKSFSETINYKQESYRTTIDRKTNHSIGWGEAHKIMNNGWGPYGRDSYDPTYGNE 217
| | | | | : : : : : | | | | | : : : : : | | : : : |
QY 195 LFMKTRNGSMKAADNFDLPNKASSLLSSGFSDFPDTATVITMDRKASKQOTNIDVIYVRD 254
| | | | | : : : : : | | | | | : : : : : | | : : : |
DB 218 LFLGRQSSNAGQNFPLTHQMLLARGNFNPEFISVLHKQNDTK-KSKIKVIYQREMD 276
| | | | | : : : : : | | | | | : : : : : | | : : : |
QY 255 DYQLHWTSTNMGKTNKDKWTDSSERYKIDWE 287
| | | | | : : : : : | | | | | : : : : : | | : : : |
DB 277 RYTQWNRHLHWGNNYKNQNTVTFSTYEVDWQ 309
| | | | | : : : : : | | | | | : : : : : | | : : : |

RESULT 5
AAU37136
ID AAU37136 standard; Protein; 327 AA.
AC AAU37136;
XX
DT 14-FEB-2002 (first entry)
DE
DE Staphylococcus aureus cellular proliferation protein #1306.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
PN
PN WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
XX
XX N-PSDB: AAS54995.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12729; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 327 AA;
XX
XX Query Match 26.7%; Score 395.5; DB 22; Length 327;
XX Best Local Similarity 30.0%; Pred. No. 1.1e-22;
XX Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKTGDLVTDKENGHKKVYFSDIDKNNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
| | | | | : : : : : | | | | | : : : : : | | : : : |
DB 43 TLYKTTATSDNDKLN-ISOILTFNFKDKSYDKDTLVLKAAGNINSYGKKPNKPDYNSQ 101
| | | | | : : : : : | | | | | : : : : : | | : : : |
QY 78 LAMPFAFKVQLQDPNEVAQISDYPRNSIDTXEMXNTYTXGXNNXNTYDXTXKXGLXG 137
| | | | | : : : : : | | | | | : : : : : | | : : : |
DB 102 FYWGGKYNVSSESNDVAVNVYAPKN--QNEEFQVQOTLGYSYG--GDINISNGLSGG 157
| | | | | : : : : : | | | | | : : : : : | | : : : |
QY 138 XN--XSXGXTXXVQDPFKTILESTPD-KKYGWKVIFNNVNNQNGPYDRDSWNPVYGNQ 194
| | | | | : : : : : | | | | | : : : : : | | : : : |
DB 158 LNSGKSFSETINYKQESYRTTIDRKTNHSIGWGEAHKIMNNGWGPYGRDSYDPTYGNE 217
| | | | | : : : : : | | | | | : : : : : | | : : : |
QY 195 LFMKTRNGSMKAADNFDLPNKASSLLSSGFSDFPDTATVITMDRKASKQOTNIDVIYVRD 254
| | | | | : : : : : | | | | | : : : : : | | : : : |
DB 218 LFLGRQSSNAGQNFPLTHQMLLARGNFNPEFISVLHKQNDTK-KSKIKVIYQREMD 276
| | | | | : : : : : | | | | | : : : : : | | : : : |
QY 255 DYQLHWTSTNMGKTNKDKWTDSSERYKIDWE 287
| | | | | : : : : : | | | | | : : : : : | | : : : |
DB 277 RYTQWNRHLHWGNNYKNQNTVTFSTYEVDWQ 309
| | | | | : : : : : | | | | | : : : : : | | : : : |

RESULT 6
AAU37564
ID AAU37564 standard; Protein; 327 AA.
AC AAU37564;
XX
XX 13-FEB-2002 (first entry)
XX
XX Staphylococcus aureus cellular proliferation protein #1734.
DE Antisense; prokaryotic cellular proliferation protein;
XX antibiotic; antibacterial; drug design.
XX Staphylococcus aureus.
XX
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX

DR WPI; 2001-611495/70.
XX N-PSDB; AAS55423.
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
PS Example 3; Seq ID No 13157; 51lpp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 327 AA;

Query Match 26.7%; Score 395.5; DB 22; Length 327;
Best Local Similarity 30.0%; Pred. No. 1.1e-22;
Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKGTGLVTVYDENGHHKVFYSFIDDKHNKKLLVIRTKGTAGQYRVYSEBANKSG 77
DB 43 TLKTTATSDNDKLN-ISQILTFNFKDKYDKDTLVKAAGNINSYGKPNPKDYNYSQ 101

QY 78 LAMPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXNTXGXNXXTXDXTKXKXGLXG 137
DB 102 FYNGGKYNVSSESDAVNVVDYAPKN--QNEFVQOQLGYSG--GDINTSGLSGG 157

QY 138 XN--XSGXYKXVQDPFKTILSPDT-KKVGWKVIFNNMVNMGVYDRDSWNPVYGNQ 194
DB 158 LNSKSFSEITNKQESYRTTIDRKTNHSKIGWGEAHTMNGWGPYGRDSDPTGYNE 217

QY 195 LEWKTRNGSKAADNFDLPNPKASSLSGSPDPFATVITMDRASKQOQTNDIYIERVRD 254
DB 218 LFLGGROSSNAGQNFILPTHOMPLLAGNFPFISVLSHKQNDTK-KSKIKVTYQREMD 276

QY 255 DYQLHWTSTNKGKNTKDKWTDSSERYKIDWE 287
DB 277 RYTNQWRLHWGNGNXXQNTVFTSTIYEDWQ 309

RESULT 7
AAU34287
ID AAU34287 standard; Protein; 325 AA.
XX
XX AAU34287;
XX
XX 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #563.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.

XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haseibeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR N-PSDB; AAS52146.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 5783; 51lpp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 325 AA;

Query Match 26.1%; Score 388; DB 22; Length 325;
Best Local Similarity 29.9%; Pred. No. 4.4e-22;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

QY 9 TGTDTI-GSNTTV--KTGDLVTVYDKENG-----MHKKVFYSFIDDKHNKKLLVIR 56
DB 21 SGTANAEGKITPVSVKVKVDDKVTLYKTATADSKFKISQILTFNFKDKSYKDXDTLVLK 80

QY 57 TKGTIAGQYRVYSEBANKSGLAMPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXNT- 115
DB 81 ATGNINSGFVKPNPDYDFSKLWYGAKNVYSISSQSDNSVNVVDYAPKNQNEEFQVQNTL 140

QY 116 -XTXGXNXXTXDXTKXKXGLXGXNXXSGXTXKXVQDPFKTILSPDT-KKVGWKVFN 173
DB 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINIKQESYRTTILSRNTNRYNMGVGEA 196

QY 174 MVNONNGPYDRDSWNPVYGNOLFMKFTNGSKMAADNFDLPNPKASSLSGSPDPFVAVIT 233
DB 197 IMNNGWGPYDRDSFHPYTYGNEFLAGQSSAYAGQNFIAHQHPLLSRSNFPNPEFLSVLS 256

QY 234 MDRKASKQOQTNDIYIERVRDDYQLHWTSTNKGKNTKDKWTDSSERYKIDWEKEM 291
DB 257 -HRQDGAKKSKIYTVYQREMDLIQIRWNGFYWAGANYKNEKTRFKSTYIEDNENHKV 313

RESULT 8
AAU37100

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:41:14 ; Search time 20.5 Seconds
(without alignments)
420.532 Million cell updates/sec

Title: CLAIM7

Perfect score: 1484

Sequence: 1 ADSDNIKTGTDIGSNTV.....WTRSSRYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	293	19.7	336	US-08-666-405-28	Sequence 28, Appl
2	133.5	9.0	10182	US-09-134-001C-3159	Sequence 3159, Ap
3	124.5	8.4	1231	US-09-071-035-420	Sequence 420, App
4	124.5	8.4	1265	US-09-071-035-418	Sequence 418, App
5	118	8.0	447	US-09-627-376-10	Sequence 10, Appl
6	118	8.0	1584	US-09-457-040B-27	Sequence 27, Appl
7	115	7.7	461	US-08-392-625-24	Sequence 24, Appl
8	115	7.7	461	US-08-466-961A-24	Sequence 24, Appl
9	113.5	7.6	246	US-08-645-193B-26	Sequence 26, Appl
10	113.5	7.6	246	US-08-276-151-7	Sequence 7, Appl
11	112	7.5	512	US-08-856-253-6	Sequence 6, Appl
12	112	7.5	1183	US-08-447-031A-2	Sequence 2, Appl
13	110.5	7.4	1094	US-09-268-347-32	Sequence 32, Appl
14	107	7.2	2314	US-09-268-347-49	Sequence 49, Appl
15	106.5	7.2	1751	US-09-136-574A-44	Sequence 44, Appl
16	104.5	7.0	624	US-08-947-965-78	Sequence 78, Appl
17	104.5	7.0	655	US-08-469-202-28	Sequence 28, Appl
18	104.5	7.0	655	US-08-484-434C-35	Sequence 35, Appl
19	104.5	7.0	679	US-08-913-942-15	Sequence 15, Appl
20	104.5	7.0	679	US-09-268-347-26	Sequence 26, Appl
21	104	7.0	464	US-09-426-072-2	Sequence 2, Appl
22	103.5	7.0	1435	US-08-568-459A-4	Sequence 4, Appl
23	103.5	7.0	1435	US-08-487-826B-4	Sequence 4, Appl
24	103.5	7.0	1435	US-09-210-288-4	Sequence 4, Appl
25	103.5	7.0	1833	US-08-621-944A-4	Sequence 4, Appl
26	103.5	7.0	1833	US-08-945-567D-4	Sequence 4, Appl
27	103.5	7.0	1992	US-08-621-944A-3	Sequence 3, Appl

28 103.5 7.0 1992 4 US-08-945-567D-3 Sequence 3, Appli
29 103.5 7.0 2048 4 US-09-268-347-48 Sequence 48, Appl
30 102.5 6.9 345 4 US-08-856-253-7 Sequence 7, Appl
31 102.5 6.9 655 1 US-08-469-202-27 Sequence 27, Appl
32 102.5 6.9 933 3 US-08-484-434C-34 Sequence 34, Appl
33 102.5 6.9 933 3 US-08-293-728-2 Sequence 2, Appl
34 102.5 6.9 933 4 US-09-421-868-2 Sequence 2, Appl
35 102.5 6.9 1287 1 US-08-200-232-2 Sequence 2, Appl
36 102.5 6.9 1287 5 PCT-US95-02219-2 Sequence 2, Appl
37 102.5 6.9 1287 5 PCT-US95-02219A-2 Sequence 2, Appl
38 102.5 6.9 1401 4 US-09-127-670-6 Sequence 6, Appl
39 102 6.9 1098 1 US-08-409-995-2 Sequence 2, Appl
40 102 6.9 1098 3 US-08-685-467-2 Sequence 2, Appl
41 102 6.9 1098 4 US-09-377-155-32 Sequence 32, Appl
42 102 6.9 1098 4 US-08-913-942-2 Sequence 2, Appl
43 102 6.9 1098 4 US-09-669-974-32 Sequence 32, Appl
44 102 6.9 1098 4 US-09-268-347-44 Sequence 44, Appl
45 101.5 6.8 600 4 US-09-388-743-22 Sequence 22, Appl

ALIGNMENTS

RESULT 1

US-08-666-405-28

; Sequence 28, Application US/08666405

; Patent No 5874220

; GENERAL INFORMATION:

; APPLICANT: FACH, Patrick; GUILLOU, Michel

; APPLICANT: Jean-Pierre; POPOFF, Michel

; TITLE OF INVENTION: PRIMERS FOR THE

; TITLE OF INVENTION: AMPLIFICATION OF GENES CODING FOR THE

; TITLE OF INVENTION: ENTEROTOXIN AND THE LEICITHINASE OF CLOSTRIDIUM

; TITLE OF INVENTION: PERRINGENS AND THEIR APPLICATION TO THE

; TITLE OF INVENTION: DETECTION AND NUMERATION OF THESE BACTERIAE

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIERMAN & MUSERLIAN

; STREET: 600 THIRD AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10016

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/666,405

; FILING DATE: 08-NOV-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP94/04292

; FILING DATE: 22-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/172,026

; FILING DATE: 22-DEC-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: MUSERLIAN, CHARLES A

; REGISTRATION NUMBER: 19,683

; REFERENCE/DOCKET NUMBER: 102.164

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 661-8000

; TELEFAX: (212) 661-8002

; INFORMATION FOR SEQ ID NO: 28:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 336 amino acids

; TYPE: amino acids

; STRANDEDNESS: single

; TOPOLOGY: linear

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

QY 191 YGNQLFMKTRNGSMKAADNFDLPNKASSLLSSGFSDFATVITMD-----RKASKQQT 243

QY 201 15INWGIN1NDKWIDKSSERIN1DWEKEEMI
:| : : | : : : :
:| : : | : : : :

Db 1024 EISSSSKIRPKYSYQKEI-----EDNRRTISGGEKKNKYI----- 1063

Qy 202 GSKAADNPLDNKASSLLSSGSPDFATVITMDRKASKQQTNDIVYERDD 255

Db 1064 -----DNQMDPHQIGSMDSDGLLPDQGGPPDEKNSKTLISNEQIRYLOQRKD 1111

RESULT 7

US-08-392-625-24

; Sequence 24, Application US/08392625

; Patent No. 5837485

; GENERAL INFORMATION:

; APPLICANT: Entian, Karl-Dieter

; APPLICANT: G tz, Friedrich

; APPLICANT: Schnell, No. 5837485bert

; APPLICANT: Augustin, Johannes

; APPLICANT: Engelke, Gernar

; APPLICANT: Rosenstein, Ralf

; APPLICANT: Kaletta, Cortina

; APPLICANT: Klein, Cora

; APPLICANT: Wieland, Bernd

; APPLICANT: Kupke, Thomas

; APPLICANT: Jung, G nther

; APPLICANT: Kellner, Roland

; TITLE OF INVENTION: Biosynthetic Process For The Preparation

; OF CHEMICAL COMPOUNDS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox

; STREET: 1100 New York Avenue

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/392,625

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/876,791

; FILING DATE: 30-APR-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Esmond, Robert W.

; REGISTRATION NUMBER: 32,893

; REFERENCE/DOCKET NUMBER: 0652.0980002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 461 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-08-392-625-24

Query Match

Best Local Similarity 7.7%; Score 115; DB 2; Length 461;

Matches 63; Conservative 57; Mismatches 118; Indels 52; Gaps 13;

Qy 4 DINKTGTDTGS-----NTTVKGTGLVTDYDKENGHKHKKFYFIDDKHNHKKLLV--IR 56

Db 149 DSGVNSHTDLKSNKINKVINEYKPKNGFGSGENSEGNK-----NFEEDKLNHGTLVAGQIG 203

Qy 57 TKGTIAG-----QYRVYSEEGANKSGLAWPSAFKVLQQLPDNEVAQIS-----DYPR 104

Db 204 ANGLKGVNFGVEMNRYRVFGSKASEMLWVSKGLIDAANDNDVINSLGNLYLIKDNQNK 263

Qy 105 NSIDTXEMXTXTGXNAXNTXDTXKXGLXGNXSGXTXKXVQP-DFKTILESPTDK 163

Db 264 KKLDRDEKVDYDALQKAINYAQ---KGSIVVAAGVNDGINVKVKKEINKRNLNSKTSK 320

Qy 164 KVGKVIFFNNVNNQN---WGPYDRDSWNPV---YGNOLF-MKTRNGSMKAADNPLDPNK 215

Db 321 KV-----YDSPANLNNVMTVGSIDNDYISEFSYGNFNFDLMTTIGGSYK----LLDKYG 371

Qy 216 ASSLLSSGSPDFATVITMDRKASKQQTNDIVYERVR-----DDYQL 258

Db 372 KDAWLEKGYMQKQSVLSTSSNGRYIYQSGTSLAAPKVS GALALEIDKYQL 421

RESULT 8

US-08-466-961A-24

; Sequence 24, Application US/08466961A

; Patent No. 5843709

; GENERAL INFORMATION:

; APPLICANT: Entian, Karl-Dieter

; APPLICANT: G tz, Friedrich

; APPLICANT: Schnell, No. 5843709bert

; APPLICANT: Augustin, Johannes

; APPLICANT: Engelke, Gernar

; APPLICANT: Rosenstein, Ralf

; APPLICANT: Kaletta, Cortina

; APPLICANT: Klein, Cora

; APPLICANT: Wieland, Bernd

; APPLICANT: Kupke, Thomas

; APPLICANT: Jung, G nther

; APPLICANT: Kellner, Roland

; TITLE OF INVENTION: Biosynthetic Process for the Preparation of

; CHEMICAL COMPOUNDS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox

; STREET: 1100 New York Avenue, NW

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/466,961A

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/392,625

; FILING DATE: 22-FEB-1995

; APPLICATION NUMBER: US 07/876,791

; FILING DATE: 30-APR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/784,234

; FILING DATE: 31-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Esmond, Robert W.

; REGISTRATION NUMBER: 32,893

; REFERENCE/DOCKET NUMBER: 0652.0980004

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 461 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-08-466-961A-24

Query Match

7.7%; Score 115; DB 2; Length 461;

149	Db	DSGVNSSHTDLKSNKIVNEVPKNGRFGSEDSGNK-----NFEEDKLNHGFLVAGQIG	203
57	QY	TGKTIAG-----QYRVTSSEAGKSLAWSAFKVQLQLPDNEVAQIS-----DYPR	104
204	Db	ANGNLKGVNPGVMNVYRVFGSKSEMLWYSGKIIDAANDNDVINVSLGNYLTKNQNK	263
105	QY	NSIDTDEXMATTXTGXNKNXNTDXTKTKKGLXGNKXSGXTKKXVP--DEFTTILESPDK	163
264	Db	KKLPDDEKVDYDALOKAINAQ--KKGSIVVAAGVNDGVNKKVKEINKRNLNSKTSK	320
164	QY	KVGWKVIFNNMVNQN-----NGPYDRDSWNVP--YGNQLF-MKTRNGSKMAANDFLDPNK	215
321	Db	KV-----YDSPALNNVMTVGSIDDDNDYISEFSNYGNFIDLMTIGSKY-----LLDKYG	371
216	QY	ASSLLSSGFSFPDFATVITMDRKASKQQTNDIVTIERVR-----DDYQL	258
372	Db	KDAWLEKGYKOKGVSLTSSNGRYIYOSGTSLAAPKVSAGALAEIDKYOKI	421

```

; Patent NO. 5397719
; GENERAL INFORMATION:
; APPLICANT: Freed, Ellen
; APPLICANT: Ruggieri, Rosamaria
; TITLE OF INVENTION: Interaction of raf-1 and 14-3-3 Proteins
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward et al.
; STREET: Five Palo Alto Square
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,151
; FILING DATE: 14-JUL-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Ph.D., Timothy E
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: ONYX-005/000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 843-5481
; TELEFAX: (415) 857-0663
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-276-151-7

```

QY	79	AWPSAFKVLQLPD-NEVAQISDY-----PRNSI-----DTXEXMTXTGXNX 122
Db	81	GKEYREKIEAELQICNDVLELLDKLIPNATQESKVFLMKMGDYFRYLSEVASDGNK 140
QY	123	NXYTDXTKFXGLXGXNNXSGTXFKXVQPDKTILSPSTDKKVGWKVFNMVMNQNGWY 182


```

; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: 1997US001/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; TELE: <Unknown>
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1751 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-136-574A-44

Query Match          7.2%; Score 106.5; DB 4; Length 1751;
Best Local Similarity 19.7%; Pred. NO. 1.4e+02;
Matches 62; Conservative 52; Mismatches 102; Indels 99; Gaps 16;

QY 27 TYDKENGHHKVFYFIDDKNHKKLLVIRKGTIAGQYRVY-SEEGANK----- 75
Db 543 TYIEGG---KISGPYVMDKNNIYVLFVSST-----KIYPGGEVEHKKQAQFKISVP 594

QY 76 SGLAW-----PSAEKVQLQDPNEVAQISDYYPNSI-----DTXE 111
Db 595 QGYPDPTNDPSYKGLTSOLEKNK--YIAAYDNNNLVWGLEPGAATSTPAPTSTPTPTPT 652

QY 112 XMXTXGXGNXNXTDXTKXGXLXG-----XNXSXGXTXKXVQPDFKTIILESPDTRKVG 166
Db 653 PPTVTATPTPTPTPTGSPGSGSVKLYKNNETSASTGSIKVPFKIVNGSS----- 707

QY 167 WKVFNMMVNQNWGPYDRD-----SNPYYGNOLFEMKTRNGSMKAADNFDLPNKASSL 219
Db 708 -SVDLSRVKIRYWTYVDGDKPQSAVCDWAI-----GASNVTFFVK----- 748

QY 220 LSSGFS-PDFATVITMDRKASKQQTNIQVIERVRRDDYQLHWTSTNWKGTNKKKWTDRS 278
Db 749 LSSGVSGADYILEVGFSSGAGLQPGKDT-----GDIQVRFNKNDWSNYNQADDWS---- 799

QY 279 SERKIDWEKEEMTN 293
Db 800 -----W-LQSMTN 806
```

Search completed: March 4, 2003, 10:46:03
Job time : 25.5 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:42:29 ; Search time 10.5 Seconds

(without alignments)

1176.748 Million cell updates/sec

Title: CLAIM7

Perfect score: 1484

Sequence: 1 ADSIDINIKTGTIGSNTTV.....WTDSSRYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	421.5	28.4	330	10	US-09-815-242-5293
2	421.5	28.4	333	10	US-09-815-242-12121
3	395.5	26.7	327	10	US-09-815-242-5901
4	395.5	26.7	327	10	US-09-815-242-12729
5	395.5	26.7	327	10	US-09-815-242-13157
6	388	26.1	325	10	US-09-815-242-5783
7	388	26.1	325	10	US-09-815-242-12693
8	388	26.1	325	10	US-09-815-242-12779
9	388	26.1	325	10	US-09-815-242-13070
10	383	25.8	323	10	US-09-815-242-5891
11	383	25.8	338	10	US-09-815-242-13102
12	253	17.0	321	10	US-09-815-242-5771
13	253	17.0	321	10	US-09-815-242-12694
14	253	17.0	321	10	US-09-815-242-12721
15	240	16.2	312	10	US-09-815-242-5267
16	132.5	8.9	1300	10	US-09-815-242-4903
17	132.5	8.9	1300	10	US-09-815-242-10906
18	119	8.0	791	9	US-10-055-364-41
19	118	8.0	447	12	US-10-047-676A-10

20	117.5	7.9	2478	10	US-09-815-242-5816
21	117.5	7.9	2478	10	US-09-815-242-12967
22	117	7.9	978	10	US-09-815-242-5456
23	117	7.9	1001	10	US-09-815-242-12686
24	114.5	7.7	2368	10	US-09-815-242-5635
25	114.5	7.7	2368	10	US-09-815-242-12389
26	112	7.5	512	10	US-09-813-820-6
27	112	7.5	1183	9	US-09-870-759-45
28	104.5	7.0	1349	10	US-09-815-242-5898
29	104.5	7.0	1349	10	US-09-815-242-13137
30	104	7.0	1827	10	US-09-819-247-2
31	103.5	7.0	467	10	US-09-912-020-381
32	103.5	7.0	1021	10	US-09-815-242-5471
33	103.5	7.0	1021	10	US-09-815-242-12544
34	103.5	7.0	1435	9	US-10-153-273-4
35	102.5	6.9	345	10	US-09-813-820-7
36	102.5	6.9	1290	10	US-09-881-752A-138
37	102	6.9	347	10	US-09-415-277A-11
38	102	6.9	1098	10	US-09-797-862-32
39	102	6.9	1805	9	US-09-820-843A-73
40	101	6.8	229	10	US-09-119-900-9
41	101	6.8	271	10	US-09-119-900-8
42	101	6.8	272	10	US-09-119-900-15
43	101	6.8	861	9	US-09-820-843A-109
44	101	6.8	2122	9	US-09-813-214A-9
45	101	6.8	3169	9	US-10-114-170-257

ALIGNMENTS

RESULT 1
US-09-815-242-5293
; Sequence 5293, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5293
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5293
Query Match 28.4%; Score 421.5; DB 10; Length 330;
Best Local Similarity 30.8%; Pred. No. 4.2e-22;


```
Db 218 LFLGGROSSNAGONFLPTHOMPLLARGNPFISVLSHKQNDTK-KSKIKVTYQREMD 276
QY 255 DYQLHWTSTNWKGTNTKDKWTRSSERYKIDWE 287
Db 277 RYTNQNRHLHWGNVNNYKNQNTVFTTSTYEVDWQ 309

RESULT 4
US-09-815-242-12729
; Sequence 12729, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12729
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12729

Query Match 26.7%; Score 395.5; DB 10; Length 327;
Best Local Similarity 30.0%; Pred. No. 2.9e-20;
Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKGTDLVTYDKENGMMHKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEGANCSG 77
Db 43 TLYKTTATSDNDKLN-ISQILTFNFIDKSDYKDTLVLAAGNINSYKPKNPKNYNSQ 101

QY 78 LAMPSAFKVLQPLDPNEVAQISDYYPNRSIDTXEXMXTXTXGXNXXNTXDXTKXGXGXG 137
Db 102 FYWGGKYNVSSESNDVAVVDYAPKN--QNEEFQVQOTLGYSG--GDINISNLSGG 157

QY 138 XN--XSXGTXKXVQPDFKTLSPD-KKVGKVFIFNNMVNONGPYPDRDSWNPYVGNQ 194
Db 158 LNSKSFSETINVKQESYRTTIDRKTNHKSIGWVEAHKIMNNGWGPYGRDSYDPTYGNE 217

QY 195 LFMKTRNGSMKADNFDLPNKASSLLSGSPDFATVIMDRKASKQOQNTNIDVIYERVD 254
Db 218 LFLGGROSSNAGONFLPTHOMPLLARGNPFISVLSHKQNDTK-KSKIKVTYQREMD 276

QY 255 DYQLHWTSTNWKGTNTKDKWTRSSERYKIDWE 287
Db 277 RYTNQNRHLHWGNVNNYKNQNTVFTTSTYEVDWQ 309

RESULT 5
US-09-815-242-12729
; Sequence 12729, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12729
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12729

Query Match 26.7%; Score 395.5; DB 10; Length 327;
Best Local Similarity 30.0%; Pred. No. 2.9e-20;
Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKGTDLVTYDKENGMMHKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEGANCSG 77
Db 43 TLYKTTATSDNDKLN-ISQILTFNFIDKSDYKDTLVLAAGNINSYKPKNPKNYNSQ 101

QY 78 LAMPSAFKVLQPLDPNEVAQISDYYPNRSIDTXEXMXTXTXGXNXXNTXDXTKXGXGXG 137
Db 102 FYWGGKYNVSSESNDVAVVDYAPKN--QNEEFQVQOTLGYSG--GDINISNLSGG 157

QY 138 XN--XSXGTXKXVQPDFKTLSPD-KKVGKVFIFNNMVNONGPYPDRDSWNPYVGNQ 194
Db 158 LNSKSFSETINVKQESYRTTIDRKTNHKSIGWVEAHKIMNNGWGPYGRDSYDPTYGNE 217

QY 195 LFMKTRNGSMKADNFDLPNKASSLLSGSPDFATVIMDRKASKQOQNTNIDVIYERVD 254
Db 218 LFLGGROSSNAGONFLPTHOMPLLARGNPFISVLSHKQNDTK-KSKIKVTYQREMD 276

QY 255 DYQLHWTSTNWKGTNTKDKWTRSSERYKIDWE 287
Db 277 RYTNQNRHLHWGNVNNYKNQNTVFTTSTYEVDWQ 309

RESULT 6
US-09-815-242-5783
; Sequence 5783, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
```

```
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5783
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5783

Query Match          26.1%; Score 388; DB 10; Length 325;
Best Local Similarity 29.9%; Pred. No. 9.8e-20;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

Qy 9 TGTDDI-GSNTTV---KTGDLVYDKENG-----MHKKVYFSDIDKHNKLLVIR 56
Db 21 SGTANAEGKITPVSVKVVDDKVTLYKTTATADSKFKISQILTFFIKDSYDKDLVLK 80
Qy 57 TKGTTAGYRVYSEEGANKSLWPAFAKVOQLPDPNEVAOISDYYPNSIDTXXMXT- 115
Db 81 ATGINSGFVFPNPNDDYDFSKLYGAKYNVSISSQSDSNVVDYAPKNQNEEFQVQNTL 140
Qy 116 -XTGXNXXNTXDXTKXGXGLXGNXSGXTXKXVQDPDKTILESPTD-KKVGKVIPIFN 173
Db 141 GYTFGGDISISNGL---SGGLNG-NTAFSEIYNYKQESYRTLSRNTNYKNVGVGEAHL 196
Qy 174 MVNQWGPYDRDSNPVYGNQLFMKTRNGSKAADNFDLPNKASSLLSGSFDPFATVIT 233
Db 197 IMNNGWGPYGRDSFHTYGNELFLAGROSSAYAGQNTAQHQMPLLSRNPFELSVLS 256
Qy 234 MDRKASKQOTNIDYIERVRDDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEEM 291
Db 257 -HRDGAKKSKITVYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 313

RESULT 7
US-09-815-242-12693
; Sequence 12693, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12693
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12693

Query Match          26.1%; Score 388; DB 10; Length 325;
Best Local Similarity 29.9%; Pred. No. 9.8e-20;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

Qy 9 TGTDDI-GSNTTV---KTGDLVYDKENG-----MHKKVYFSDIDKHNKLLVIR 56
Db 21 SGTANAEGKITPVSVKVVDDKVTLYKTTATADSKFKISQILTFFIKDSYDKDLVLK 80
Qy 57 TKGTTAGYRVYSEEGANKSLWPAFAKVOQLPDPNEVAOISDYYPNSIDTXXMXT- 115
Db 81 ATGINSGFVFPNPNDDYDFSKLYGAKYNVSISSQSDSNVVDYAPKNQNEEFQVQNTL 140
Qy 116 -XTGXNXXNTXDXTKXGXGLXGNXSGXTXKXVQDPDKTILESPTD-KKVGKVIPIFN 173
Db 141 GYTFGGDISISNGL---SGGLNG-NTAFSEIYNYKQESYRTLSRNTNYKNVGVGEAHL 196
Qy 174 MVNQWGPYDRDSNPVYGNQLFMKTRNGSKAADNFDLPNKASSLLSGSFDPFATVIT 233
Db 197 IMNNGWGPYGRDSFHTYGNELFLAGROSSAYAGQNTAQHQMPLLSRNPFELSVLS 256
Qy 234 MDRKASKQOTNIDYIERVRDDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEEM 291
Db 257 -HRDGAKKSKITVYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 313

RESULT 8
US-09-815-242-12779
; Sequence 12779, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; CURRENT APPLICATION NUMBER: 60/242,578
```


;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 12779
;; LENGTH: 325
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-09-815-242-12779

Query Match 26.1%; Score 388; DB 10; Length 325;
Best Local Similarity 29.9%; Pred. No. 9.8e-20;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

QY 9 TGTDTI-GSNTTV---KTGDLVTVYDKENG-----MHKKVYFSFIDDKNNHKKLLVIR 56
Db 21 SGTANAEGKITPVSVKVKDDKVTLYKTATADSDKFKISQILITFNFKDKSYDKDTLVLK 80

QY 57 TGTIAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTXEXMXT- 115
Db 81 ATGINSFGVKPNPDYDFSKLYWGAKYNYVSISSQSDNSVNVVDYAPKQNEEFQVQNTL 140

QY 116 -XTGXNKNXTDXTKXKXGLXGNXSGXTXKXVQPDFKTILESPTD-KKVGHKVIFNN 173
Db 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINYKQESYRTLSRNTNYKNVGGVGAHK 196

QY 174 MVNQNWGPYDRDSNPNVYGNQFMKTRNGSKMAADNFDPNKAASSLLSSGFSDFATVIT 233
Db 197 IMNNGWGPYGRDSFHPYTYGNEFLAGQSSAYAGQNFIAQHOMPLLSRSNFNPEFLSVLS 256

QY 234 MDRKASQOQNIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDSSRYKIDKEKEM 291
Db 257 -HRQDGAKKSKITVYQREMDLYQIRWNGFWAGANYKNEKTRTFKSTYEIDWENHKV 313

RESULT 9
US-09-815-242-13070
;; Sequence 13070, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Kari L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011a
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110

;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13070
;; LENGTH: 325
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-09-815-242-13070

Query Match 26.1%; Score 388; DB 10; Length 325;
Best Local Similarity 29.9%; Pred. No. 9.8e-20;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

QY 9 TGTDTI-GSNTTV---KTGDLVTVYDKENG-----MHKKVYFSFIDDKNNHKKLLVIR 56
Db 21 SGTANAEGKITPVSVKVKDDKVTLYKTATADSDKFKISQILITFNFKDKSYDKDTLVLK 80

QY 57 TGTIAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTXEXMXT- 115
Db 81 ATGINSFGVKPNPDYDFSKLYWGAKYNYVSISSQSDNSVNVVDYAPKQNEEFQVQNTL 140

QY 116 -XTGXNKNXTDXTKXKXGLXGNXSGXTXKXVQPDFKTILESPTD-KKVGHKVIFNN 173
Db 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINYKQESYRTLSRNTNYKNVGGVGAHK 196

QY 174 MVNQNWGPYDRDSNPNVYGNQFMKTRNGSKMAADNFDPNKAASSLLSSGFSDFATVIT 233
Db 197 IMNNGWGPYGRDSFHPYTYGNEFLAGQSSAYAGQNFIAQHOMPLLSRSNFNPEFLSVLS 256

QY 234 MDRKASQOQNIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDSSRYKIDKEKEM 291
Db 257 -HRQDGAKKSKITVYQREMDLYQIRWNGFWAGANYKNEKTRTFKSTYEIDWENHKV 313

RESULT 10
US-09-815-242-5891
;; Sequence 5891, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Kari L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011a
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5891
;; LENGTH: 323
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-09-815-242-5891

Query Match 25.8%; Score 383; DB 10; Length 323;

```
Best Local Similarity 28.6%; Pred. No. 2.2e-19;
Matches 82; Conservative 79; Mismatches 108; Indels 18; Gaps 8;

QY 13 DIGSNTTGTGLVTDKENGHHKVFYFIDDKNNHKKLLVIRKGTIAGQYRVYSEEG 72
Db 36 NLDGDTKMTYRATTSDSQNKNTQSQFNFLEPNYDKETVFIKAGTTIGSLRLDPNG 95

QY 73 ANKSLGAWPSAFKVLQ-LPDNEVAQISDYYPNRNSIDTXEXMXTXTXGXNXXNXTDXTKK 131
Db 96 YWNSTLRWPGSYSVSIQNVDNNNTNVTDFAPKNODESREVYK----YGYKTGGDFSN 151

QY 132 XGXLG---XNXSGXTXKXVQDPDKTILESTPTDKK-VGWKV---IFNNMVNQNGPYDR 184
Db 152 RGGLTGNITKESYSETISYQSPSYRTLLDQSTSHKGVGWKEVAHLINNMGHDTROLTN 211

QY 185 DSWNPVYGQLPMKTRNGSMKAADNFPDNKASSLLSSGFSDFATVITMDRKASKQOQTN 244
Db 212 DSDNRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTVSEGFNPEFLAVMWSHDKK-DGKRSQ 269

QY 245 IDVIYERDDYQLHWTSTN----WKGNTNTKDKWTRDRSSERYKIDWE 287
Db 270 FVHYKRSMDKIDNRRHGFNGWNGENHVDKKEKLSALYVDWK 316

RESULT 11
US-09-815-242-13102
; Sequence 13102, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13102
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13102

Query Match 25.8%; Score 383; DB 10; Length 338;
Best Local Similarity 28.6%; Pred. No. 2.4e-19;
Matches 82; Conservative 79; Mismatches 108; Indels 18; Gaps 8;

QY 13 DIGSNTTGTGLVTDKENGHHKVFYFIDDKNNHKKLLVIRKGTIAGQYRVYSEEG 72
Db 42 NLDGDTKMTYRATTSDSQNKNTQSQFNFLEPNYDKETVFIKAGTTIGSLRLDPNG 101

QY 73 ANKSLGAWPSAFKVLQ-LPDNEVAQISDYYPNRNSIDTXEXMXTXTXGXNXXNXTDXTKK 131
```

```
Db 102 YWNSTLRWPGSYSVSIQNVDNNNTNVTDFAPKNODESREVYK----YGYKTGGDFSN 157
QY 132 XGXLG---XNXSGXTXKXVQDPDKTILESTPTDKK-VGWKV---IFNNMVNQNGPYDR 184
Db 158 RGGLTGNITKESYSETISYQSPSYRTLLDQSTSHKGVGWKEVAHLINNMGHDTROLTN 217

QY 185 DSWNPVYGQLPMKTRNGSMKAADNFPDNKASSLLSSGFSDFATVITMDRKASKQOQTN 244
Db 218 DSDNRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTVSEGFNPEFLAVMWSHDKK-DGKRSQ 275

QY 245 IDVIYERDDYQLHWTSTN----WKGNTNTKDKWTRDRSSERYKIDWE 287
Db 276 FVHYKRSMDKIDNRRHGFNGWNGENHVDKKEKLSALYVDWK 322

RESULT 12
US-09-815-242-5771
; Sequence 5771, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5771
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5771

Query Match 17.0%; Score 253; DB 10; Length 321;
Best Local Similarity 24.9%; Pred. No. 3.6e-10;
Matches 71; Conservative 72; Mismatches 118; Indels 24; Gaps 12;

QY 13 DIGSNTTV--KTGDLVTDKENGHHKVFYFIDDKNNHKKLLVIRKGTIAGQYRVYSE 70
Db 47 DIGQGAETIKRTQDITS--KRLAITQNIQDFEVKDKKYNKDALVVKMOGFSSR-TTYS 103

QY 71 --EGANKSLGAWPSAFKVLQ-LPDNEVAQISDYYPNRNSIDTXEXMXTXTXGXNXXNXTDX 128
Db 104 LKYPYIKRMWIPQYINISLTKDSNV-DLINYLPKNKIDSAD--VSQKLGYNIGGNFQS 160

QY 129 TXKYGXGXLGXNKSXCTXKXVQDPDKTILESTPTDKKVGWKEVAHLINNMGHDTROLTN 188
Db 161 APSIGGSGSFNYS--KTISYNQKNYFTEVSQNSKGVKNGVANSFVTPN-----GV 211

QY 189 PVYGNQLPMKTRNGSMKAADNFPDNKASSLLSSGFSDFATVITMDR-KASKQOQTNIDV 247
```

Db 212 SAYDQYLFAOPTGP-AARDYFVDPNQLPPLIOGFPNPSFITTLSHERGKDKSE--FEI 268
QY 248 IYERVRDYLHWTSTNNKGTNTKDKWTDRS-SERYKIDWEKEEM 291
Db 269 TYGRNMDATYAYVTRHRLAVDRKHDAPKRNVTVKYEVNWKTHEV 313

RESULT 13
US-09-815-242-12694
; Sequence 12694, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12694
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12694

Query Match 17.0%; Score 253; DB 10; Length 321;
Best Local Similarity 24.9%; Pred. No. 3.6e-10;
Matches 71; Conservative 72; Mismatches 118; Indels 24; Gaps 12;
QY 13 DIGSNTTV--KTGDLVTYDKENGHMKVYFSDDKNNHKKLLVIRTKGTIAGQYRVYSE 70
Db 47 DIGOGAEIIRKTQDITS--KRLAITQNIQDFVKKDKYKNDALVVKMGQFISSR-TTYS 103
QY 71 --EGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTXEMXMTXTXGXNKNXTDX 128
Db 104 LKXYPYIKRMWPFQYNISLTKTDSNV-DLNLPLPKNKIDSAD--VSOKLGYNIGGNFQS 160
QY 129 TXKXGXLXGNXSGXTXKXVQDPFKTILESPTDKKVGKVIFFNMVQNWNGPYDRDSWN 188
Db 161 APSIGSGSFNYS--KTIISYQKNYVTEVESQNSKGKVGKANSFVTPN-----GQV 211
QY 189 PVYGNQLFMKTRNGSKMAADNFDLPNKASSLSSGSPDFATVITMDR-KASKQQTNDV 247
Db 212 SAYDQYLFAOPTGP-AARDYFVDPNQLPPLIOGFPNPSFITTLSHERGKDKSE--FEI 268
QY 248 IYERVRDYLHWTSTNNKGTNTKDKWTDRS-SERYKIDWEKEEM 291
Db 269 TYGRNMDATYAYVTRHRLAVDRKHDAPKRNVTVKYEVNWKTHEV 313

RESULT 14
US-09-815-242-12721
; Sequence 12721, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12721
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12721

Query Match 17.0%; Score 253; DB 10; Length 321;
Best Local Similarity 24.9%; Pred. No. 3.6e-10;
Matches 71; Conservative 72; Mismatches 118; Indels 24; Gaps 12;
QY 13 DIGSNTTV--KTGDLVTYDKENGHMKVYFSDDKNNHKKLLVIRTKGTIAGQYRVYSE 70
Db 47 DIGOGAEIIRKTQDITS--KRLAITQNIQDFVKKDKYKNDALVVKMGQFISSR-TTYS 103
QY 71 --EGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTXEMXMTXTXGXNKNXTDX 128
Db 104 LKXYPYIKRMWPFQYNISLTKTDSNV-DLNLPLPKNKIDSAD--VSOKLGYNIGGNFQS 160
QY 129 TXKXGXLXGNXSGXTXKXVQDPFKTILESPTDKKVGKVIFFNMVQNWNGPYDRDSWN 188
Db 161 APSIGSGSFNYS--KTIISYQKNYVTEVESQNSKGKVGKANSFVTPN-----GQV 211
QY 189 PVYGNQLFMKTRNGSKMAADNFDLPNKASSLSSGSPDFATVITMDR-KASKQQTNDV 247
Db 212 SAYDQYLFAOPTGP-AARDYFVDPNQLPPLIOGFPNPSFITTLSHERGKDKSE--FEI 268
QY 248 IYERVRDYLHWTSTNNKGTNTKDKWTDRS-SERYKIDWEKEEM 291
Db 269 TYGRNMDATYAYVTRHRLAVDRKHDAPKRNVTVKYEVNWKTHEV 313

RESULT 15
US-09-815-242-5267
; Sequence 5267, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.

Search completed: March 4, 2003, 10:46:24
Job time : 11.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:40:49 ; Search time 27.5 Seconds
(without alignments)
1024.269 Million cell updates/sec

Title: CLAIM7
 Perfect score: 1484
 Sequence: 1 ADSGINIKCTTDIGSNVTV.....WTDRSSRYKIDWEKEMTN 293

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

```
Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 2832224
```

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database :      PIR_73:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			ID	Description
	Score	Match	Length		
1	1484	100.0	319	2 S69209	alpha-toxin precu
2	1480	99.7	319	2 F89887	Alpha-Hemolysin pr
3	397	26.8	325	2 T00161	leukocidin chain F
4	394.5	26.6	327	2 B89688	leukotoxin, Lukd [
5	394	26.5	325	2 S32212	leucocidin chain F
6	393	26.5	325	2 B49234	leucocidin R, comp
7	388	26.1	300	2 B49238	gamma-hemolysin ga
8	388	26.1	325	2 F90043	gamma-hemolysin co
9	385	25.9	338	2 C89991	hypothetical prote
10	384.5	25.9	326	2 S68224	synergohemotrophi
11	374	25.2	322	2 JCS469	Panton-Valentine l
12	374	25.2	323	2 JQ1530	leukocidin chain F
13	374	25.2	323	2 JN0627	leukocidin chain F
14	293	19.7	336	2 I40826	beta-toxin - Clost
15	262	17.7	308	2 JCS468	leukocidin chain l
16	254.5	17.1	351	2 D89991	hypothetical prote
17	253	17.0	309	2 D90043	gamma-hemolysin ch
18	252	17.0	309	2 JC4282	gamma-hemolysin II
19	248	16.7	286	2 C49238	gamma-hemolysin co
20	248	16.7	315	2 PC4078	hlgc-like protein
21	248	16.7	315	2 E90043	gamma-hemolysin co
22	247.5	16.7	311	2 C89968	leukotoxin Luke [i
23	246	16.6	321	2 S49271	hlga-like protein
24	245	16.5	315	2 A49234	leucocidin R S com
25	242	16.3	315	2 JN0626	leukocidin chain S
26	237	16.0	312	2 T00160	leukocidin chain S
27	233	15.7	312	2 S32211	leucocidin chain S
28	203.5	13.7	310	2 S68225	synergohemotrophi
29	129	8.7	1272	2 C90593	hypothetical prote

ALIGNMENTS

```

RESULT 1
S69209
  alpha-toxin precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000
C:Accession: S69209
R:Hedengr hn, G.
  submitted to the EMBL Data Library, October 1992
A:Reference number: S69209
A:Accession: S69209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <RED>
A:Cross-references: EMBL:X01645; NID:g46763; PIDN:CAA25801.1; PID:g46765
C:Superfamily: leukocidin
C:Keywords: toxin
F:1-26/Domain: signal
F:27-319/Product: alpha-toxin
#status predicted <SIG>
#status predicted <MAT>

```

```
Query Match      100.0%; Score 1484; DB 2; Length 319;
Best Local Similarity 93.2%; Pred. No. 5.8e-84;
Matches 273; Conservative 20; Mismatches 0; Indels 0; Gaps 0;
```

1	Qy	ADSDINIKTGTDTIGSNNTVKTKDGLVITYDKENGSMHKKVYFSFIDDKXNNHKKLLVIRTKGT	60
27	Db	ADSDINIKTGTDTIGSNNTVKTKDGLVITYDKENGSMHKKVYFSFIDDKXNNHKKLLVIRTKGT	86
61	Qy	IAGQRYVYSEEGANKSGLAPSAFKVQLQLPDNEVAQISDYIPRNSIDITPEXMTXTXGX	120
87	Db	IAGQRYVYSEEGANKSGLAPSAFKVQLQLPDNEVAQISDYIPRNSIDITPEXMTXTXGX	146
121	Qy	NXNKTDXTKXGXKXNKSXGTXKXXVQDPFKTILESPTDKKVGWKVIFNNMVNQNG	180
147	Db	NGNVTGDDTKIGLIGLIGANYSIGHTILKYQDPFKTILESPTDKKVGWKVIFNNMVNQNG	206
181	Qy	PYDRDSNPVYGQOLFPMKTRNGSKAADNFDLPNKASSLLSSGFSDFEATVITMDRKASK	240
207	Db	PYDRDSNPVYGQOLFPMKTRNGSKAADNFDLPNKASSLLSSGFSDFEATVITMDRKASK	266
241	Qy	QQTNDIVYIERVDDYQLHWTSTNWKGTNTKDWTDSSERYKIDWEKEEMTN	293
267	Db	QQTNDIVYIERVDDYQLHWTSTNWKGTNTKDWTDSSERYKIDWEKEEMTN	319

RESULT 2
F89887
Alpha-Hemolysin precursor [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 01-Mar-2002
C:Accession: F89887
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.;
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; S

C.: Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:213111952; PMID:11418146
A:Accession: F89887
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <KUR>
A:Cross-references: GB:BA000018; PID:g13700962; PIDN:BA42258.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1007
C:Superfamily: leukocidin

Query Match 99.7%; Score 1480; DB 2; Length 319;
Best Local Similarity 92.8%; Pred. No. 1e-83;
Matches 272; Conservative 21; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADSINIKTGTIGSNTTGTGDLVTDKNGMHKKVYFSDIDKNNKKLLVIRTKGT 60
|||||
Db 27 ADSINIKTGTIGSNTTGTGDLVTDKNGMHKKVYFSDIDKNNKKLLVIRTKGT 86
|||||
QY 61 IAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXXMXTTXGX 120
|||||
Db 87 IAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXXMXTTXGX 146
|||||
QY 121 NAXTXDXTXKXGLXGXNXXGXTXKXVQDPDKFVLESDYPRNSIDTXXMXTTXGX 180
|||||
Db 147 NGVVGDDTGKIGLIGANVSIGHLKYVQDPDKFVLESDYPRNSIDTXXMXTTXGX 206
|||||
QY 181 PYDRSWNPVYGNQLFMTKTRNGSKAADNFDLPNKASSLSGFSFDPFATVITMDRKASK 240
|||||
Db 207 PYDRSWNPVYGNQLFMTKTRNGSKAADNFDLPNKASSLSGFSFDPFATVITMDRKASK 266
|||||
QY 241 QOTNIDVYERDDYQLHWTSTNNKGTNTKDKWTDSSERYKIDWEKEEWN 293
|||||
Db 267 QOTNIDVYERDDYQLHWTSTNNKGTNTKDKWTDSSERYKIDWEKEEWN 319
|||||

RESULT 3
T00161
leukocidin chain F precursor - *Staphylococcus aureus* phase phi PVL
N:Alternate names: protein 28
C:Species: *Staphylococcus aureus* phase phi PVL
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C:Accession: T00161
R:Kaneko, J.; Kimura, T.; Kawakami, Y.; Tomita, T.; Kamio, Y.
A:Title: Pantone-Valentine leukocidin genes in a phage-like particle isolated from mitomy
A:Reference number: 214119; MUID:98067870; PMID:9404084
A:Accession: T00161
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <KAN>
A:Cross-references: EMBL:AB009866; NID:g3341907; PIDN:BA31901.1; PID:g3341935
C:Genetics:
A:Gene: lukF
C:Function:
A:Description: lukF and lufs cooperatively and strongly lyse rabbit erythrocytes besides
C:Superfamily: leukocidin
C:Keywords: hemolysis; toxin

Query Match 26.8%; Score 397; DB 2; Length 325;
Best Local Similarity 28.6%; Pred. No. 7.9e-17;
Matches 80; Conservative 72; Mismatches 108; Indels 20; Gaps 6;

QY 11 TTDIGSNTTGTGDLVTDKNGMHKKVYFSDIDKNNKKLLVIRTKGTAGQYRVYSE 70
|||||
Db 45 TTATSDSKLISQILT-----FNFIDKSDYKDTLLKKAAGNYSYGTENP 92
|||||
QY 71 EGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXXMXTTXGXNXXTXDXTX 130
|||||
Db 93 KDTISSQFYWSKYNISNSDSNVVVDYAPKN--QNEEFQVQTVGYSG--GDINI 148
|||||

QY 131 KXGLXGXN--XSXGTXKXVQDPDKFVLESDYPRNSIDTXXMXTTXGXNXXTXDXTX 187
|||||
Db 149 SNGLSGGNGSKSFSETINYKQESYRTSLDKRTNFKKIGWDVEAHKIMNNGWPGYGRDSY 208
|||||
QY 188 NPVYGNQLFMTKTRNGSKAADNFDLPNKASSLSGFSFDPFATVITMDRKASKQOTNIDV 247
|||||
Db 209 HSTYGNEMFLGSRQNLNAGQNFLEYHKMPVLSRGNFPEFVLSRQNAAK-KSKITV 267
|||||
QY 248 IYERYDDYQLHWTSTNNKGTNTKDKWTDSSERYKIDWE 287
|||||
Db 268 TYQREMDRYTNFNLHWGNIGNYKDNENRATHTSIYVDWE 307
|||||

RESULT 4
B89968
leukotoxin, LukD [pathogenicity island Sapin3] [imported] - *Staphylococcus aureus* (st
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89968
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O.
ma, A.; Mizutani-Ti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:213111952; PMID:11418146
A:Accession: B89968
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <KUR>
A:Cross-references: GB:BA000018; PID:g13701612; PIDN:BA42905.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: lukD
C:Superfamily: leukocidin

Query Match 26.6%; Score 394.5; DB 2; Length 327;
Best Local Similarity 30.0%; Pred. No. 1.1e-16;
Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKGTGLVTDKNGMHKKVYFSDIDKNNKKLLVIRTKGTAGQYRVYSEEGANKSG 77
|||||
Db 43 TLYKTATSDNDKLN-ISQILTFNFKDKSYDKDTFLVLAAGNINSYKKNPKDYNYSQ 101
|||||
QY 78 LAMPFAFVQLQLPDNEVAQISDYPRNSIDTXXMXTTXGXNXXTXDXTXKXGLXG 137
|||||
Db 102 FYWGKYNVSVSSSDNADVVDYAPKN--QNEEFQVQTVGLYSVG--GDINISNGLSG 157
|||||
QY 138 XN--XSXGTXKXVQDPDKFVLESDYPRNSIDTXXMXTTXGXNXXTXDXTXKXGLXG 194
|||||
Db 158 LNSKSFSETINYKQESYRTTIDRKNHKSIGWGVGAHKNNGWGPYGRDSYDPTYGNE 217
|||||
QY 195 LFMKTRNGSKAADNFDLPNKASSLSGFSFDPFATVITMDRKASKQOTNIDVYERVD 254
|||||
Db 218 LFLGGROSSNAGQNFVLPHTHOMPLARGNFPFVLSLHSHKNDTK-KSKIKVTVQREMD 276
|||||
QY 255 DYQLHWTSTNNKGTNTKDKWTDSSERYKIDWE 287
|||||
Db 277 RYTNQWNLHWGNIGNYKKNQNTVTFSTIYVDWQ 309
|||||

RESULT 5
S32212
leucocidin chain F - *Staphylococcus aureus*
C:Species: *Staphylococcus aureus*
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Mar-2000
C:Accession: S32212
R:Prevost, G.; Supersac, G.; Piemont, Y.
submitted to the EMBL Data Library, March 1993
A:Description: *Staphylococcus aureus* encodes two types of synergohemotrophic toxins:
A:Reference number: S32211
A:Accession: S32212
A>Status: preliminary

Query Match 26.5%; Score 393; DB 2; Length 325;
Best Local Similarity 29.2%; Pred. No. 1.4e-16;

Query Match	26.53	Score 393	DB 2	Length 325
Best Local Similarity	29.23	Pred. No. 1.4e-16		

RESULT 8
F90043
gamma-hemolysin component B [imported] - Staphylococcus aureus (strain N31)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: F90043
R:Kuroda, M.; Ohta, T.; Nchiyama, I.; Baba, T.; Yuzawa, H.; Kobavashi, I. I.

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90043
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <KUR>
A:Cross-references: GB:BA000018; PID:g13702370; PIDN:BA843511.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: hlgB
C:Superfamily: leukocidin

Query Match 26.1%; Score 388; DB 2; Length 325;
Best Local Similarity 29.9%; Pred. No. 2.8e-16;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;
QY 9 TGTDTI-GSNTTV---KTGDLVTYKENG-----MHKKVFYSFIDDKNNHKKLLVIR 56
Db 21 SGTANAEGKITPVSVKKVDDKLYKTATADSDKFKISQILTFNFKDSYDKDLVLK 80
QY 57 TKTIAQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXT- 115
Db 81 ATGNINSGFVKPNPDYDFSKLYGAKYNVSISSQSDNSVNVVDYAPKQNEEFQVQNTL 140
QY 116 -XTXGKNNKXTDXTKXGXLGXNXXGXTXKXVQDFKTTILESPD-KKVGWKFYNN 173
Db 141 GYFSGDISINGL---SGGLNG-NATFSETINIKQESTRTLSRNTNTKNGVGEAHK 196
QY 174 MVNQNNGPYDRDSWNPVYGNQLFMTKTRNGSMKAADNLFDPNKASSLLSGFSPDFATVIT 233
Db 197 IMNNGWPGYGRDSFHTYGNELFLAGRQSSAVAGQNFIAHQHPLLSRNFENPEFLSVLS 256
QY 234 MDRKASKQNTNIDVIYERVDDYQLHWTSTNWKGTNTKKTWDRSSERYKIDWEKEEM 291
Db 257 -HQDGAKKSKITVTYQREMDLQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 313

RESULT 9
C89991
Hypothetical protein SA1812 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89991
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89991
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-338 <KUR>
A:Cross-references: GB:BA000018; PID:g13701799; PIDN:BA843092.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1812
C:Superfamily: leukocidin

Query Match 25.9%; Score 385; DB 2; Length 338;
Best Local Similarity 28.2%; Pred. No. 4.7e-16;
Matches 82; Conservative 81; Mismatches 110; Indels 18; Gaps 8;
QY 13 DLSNTVTKTGLVYDKENGMMHKKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEG 72
Db 42 NLDGDTKMYTRTATTSQKNITQSLQFNFLTEPNYDKETVPIKAKGTIGSLRILDENG 101
QY 73 ANKSGLAWPSAFKVLQ-LPDNEVAQISDYPRNSIDTXEMXTYTXGNXNXTYDXTYK 131
Db 102 YNNTSLRWPGSYSVSIQNVDDNNNTVTDFAPKNODESREVKYT---YGYTGGDFDSIN 157

QY 132 XGXLXG---XNXSGXTXKXVQDFKTTILESPDKK-VGMKV---IFNNMNVQNMGPYDR 184
Db 158 RGLGNTIKENSETISYQPSRYTLDDQSTSHKGWKGVEAHLINMNGHDHTRQLTN 217
QY 185 DSWNPVYGNQLFMTKTRNGSMKAADNLFDPNKASSLLSGFSPDFATVITMDRKASKQQTN 244
Db 218 DSDNRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTVSEGFNPEFLAVMSHDKK-DKGKQ 275
QY 245 IDVIYERVDDYQLHWTSTN---WKGTNTKKTWDRSSERYKIDWEKEEM 291
Db 276 FVHYKRSMDKIDNRHGFYSGENHVDKKEKLSALYEDVWKTHDV 326

RESULT 10
S88224
synergohymenotropic toxin component F precursor - Staphylococcus intermedius
C:Species: Staphylococcus intermedius
C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000
C:Accession: S68224; S78213; S44945
R:Prevost, G.; Bouakham, T.; Piemont, Y.; Montell, H.
FEBS Lett. 376, 135-140, 1995
A:Title: Characterisation of a synergohymenotropic toxin produced by Staphylococcus
A:Reference number: S68224; MUID:96105366; PMID:7498527
A:Accession: S68224
A:Molecule type: DNA
A:Residues: 1-326 <PRE>
A:Cross-references: EMBL:X79188; NID:9487685; PIDN:CAA55783.1; PID:9487687
A:Experimental source: ATCC 51874
A>Note: the authors translated the codon TAT for residue 103 as Ile
A:Accession: S78213
A:Molecule type: protein
A:Residues: 27-43 <PRE>
A:Experimental source: ATCC 51874
C:Genetics:
A:Gene: lukF-I
C:Superfamily: leukocidin
C:Keywords: toxin
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-326/Product: synergohymenotropic toxin component F #status experimental <MAT>

Query Match 25.9%; Score 384.5; DB 2; Length 326;
Best Local Similarity 30.1%; Pred. No. 4.7e-16;
Matches 83; Conservative 72; Mismatches 108; Indels 13; Gaps 9;
QY 18 TTVKGTGLVTDYKEN-GMHKKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEGANKS 76
Db 43 TLYKT--TATADSDNLNISQLTFNFKDSYDKDTLVKKAAGNINSGTSPNPNDIYS 100
QY 77 GLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXTYTXGNXNXTYDXTKXGXLX 136
Db 101 SFYWGAKYNVSISAESKAVNVVDYAPKN--QNEEFQVQNTLGYSGF--GDISISKGLSG 156
QY 137 GKN--XSXGXTXKXVQDFKTTILESPDKK-VGMKVIFNNMNVQNMGPYDRDSWNPVYGN 193
Db 157 GLNGSEFSETINIKQESTRTIDKTDNKTIGWGEAHLINMNGAGPYGRDSFHDLYGN 216
QY 194 QLEFKTRNGSMKAADNLFDPNKASSLLSG-FSPDFATVITMDRKASKQNTNIDVIYERV 252
Db 217 ELFGGROSKINAGQNF-L-PTROMPLLAGNFPNPEFLSVLSHKPNGAK-TSKIKVITYORE 274
QY 253 RDDYQLHWTSTNWKGTNTKKTWDRSSERYKIDWEK 288
Db 275 MDEYTNWNGFHWGCTNYKNQNNATFTSFYEIDWDQ 310

RESULT 11
JC5469
Panton-Valentine leukocidin LukF-pv chain precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 20-Jun-2000
C:Accession: JC5469; PC4337
R:Kaneko, J.; Muramoto, K.; Kamio, Y.

A;Residues: 27-85 <KAM>
A;Experimental source: strain No.4 (RIMD 310925)
C;Comment: This leukocidin acts synergistically to induce cytotoxic changes in human C;Genetics:
A;Gene: lukF
C;Complex: heterodimer; chains S and F
C;Superfamily: leukocidin
C;Keywords: hemolysis; toxin
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-323/Product: leukocidin chain F #status experimental <LEU>

Query Match 25.2%; Score 374; DB 2; Length 323;
Best Local Similarity 29.6%; Pred. No. 2.1e-15;
Matches 88; Conservative 72; Mismatches 117; Indels 20; Gaps 8;

QY 9 TGTDDI-GSNTTV---KTGDLVTYDKENG-----MHKKVYFSFIDDKNHNKLLVIR 56
DB 21 SGTANAEGKIIPVSVKKVDDKVTLYKTTATADSDKFISQILTFEIKDSYDKDPLVLK 80
QY 57 TKGTIAGQYRVYSEBANKSGLANPSAKVQLQLPDNEVAQSIDYPRNSIDTXYEXMT- 115
DB 81 ATGINSNGFVKPNPDNYDFSKLYGKAYNVSISQSSDVSNAVDYAPKNQNEEFQYQNTL 140
QY 116 -XTGXNKNXTDXTKXGXLGXNKSXGXTKXKXQPOFKTILEPTDKKVGKVFNNM 174
DB 141 GYTFGGDISISNGL---SGGLNG-NTAGSETINYQESTRTLSRNTNYKNVCGWGAHAKI 196
QY 175 VNQNWGYDRDSWNPVYGNQLFMKTRNGSMKAADNLPDNPKASSLLSGFSDFATVITM 234
DB 197 MN-GWGPYGRDSFPTYGNELFLAGRQSSAVAGQNFIAQHOMPLLSSRSNPNRFLSVLSH 255
QY 235 DRKASKQOTNDIVYERVDDYQLHWTSTNKKGTNTKKWTDRSSERYKIDHEKEEM 291
DB 256 RQDAAK-KSKITVTYQREMDLYQIRWNGFYWAGANYKNFKTRFKTYEIDWENHKV 311

RESULT 13
JN0627
leukocidin chain F precursor - Staphylococcus aureus
C;Species: Staphylococcus aureus
C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-2000
C;Accession: JN0627; S32416
R;Rahman, A.; Izaki, K.; Kamio, Y.
Biosci. Biotechnol. Biochem. 57, 1234-1236, 1993
A;Title: Gamma-hemolysin genes in the same family with LukF and LukS genes in methicillin-resistant strains of Staphylococcus aureus
A;Reference number: JN0625; MUID:93372503; PMID:7763998
A;Accession: JN0627
A;Molecule type: DNA
A;Residues: 1-323 <RAH>
A;Cross-references: GB:565052; NID:g410004; PIDN:AAC60446.1; PID:g410007
R;Kamio, Y.; Rahman, A.; Nariya, H.; Ozawa, T.; Izaki, K.
FEBS Lett. 321, 15-18, 1993
A;Title: The two staphylococcal bi-component toxins, leukocidin and gamma-hemolysin,
A;Reference number: S32416; MUID:93223836; PMID:8467905
A;Accession: S32416
A;Status: preliminary
A;Molecule type: protein
A;Residues: 27-85 <KAM>
C;Genetics:
A;Gene: lukF
C;Superfamily: leukocidin
C;Keywords: toxin
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-323/Product: leukocidin chain F #status predicted <MAT>

Query Match 25.2%; Score 374; DB 2; Length 323;
Best Local Similarity 29.3%; Pred. No. 2.1e-15;
Matches 87; Conservative 73; Mismatches 117; Indels 20; Gaps 8;

QY 9 TGTDDI-GSNTTV---KTGDLVTYDKENG-----MHKKVYFSFIDDKNHNKLLVIR 56
DB 21 SGTANAEGKIIPVSVKKVDDKVTLYKTTATADSDKFISQILTFEIKDSYDKDPLVLK 80

QY 57 TKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDYYPNRSIDTXEXMXT- 115
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 81 ATGNINSGFVKPNPDYDFSKLYGAKYNYSSQSDNSNAVDYAPKQNEEFQVQNTL 140
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 116 -ATXGXNXXYDXTKXGXLGXNKSXGTYKXVQPFKFTILESPTDKKVGWKFVFNMM 174
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINYKQESRYLRSRNTYKNGWGVEAHKI 196
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 175 VNQNGPPYDRDSWNPVYGNQLPMKTRNGSMKAADNPLDPNKAASSLLSSGFSPPDPAIVTJM 234
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 197 MN-GNGPGRDSFHTYGNELFLAGRQSSAVAGQNFIAHQHPLLSRNFNPEFLSVLS- 254
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 235 DRKASKQQTNDIVYERVDYQLHWTSTNMKGTNTKDKWDRSSERYKIDWEKEM 291
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 255 HRQDRAKSKITVTYQREMDLYQIRWNGFYWAGANYKFKTKSTVEIDWENHKV 311
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
RESULT 14
140826
beta-toxin - Clostridium perfringens
C:Species: Clostridium perfringens
C:Date: 16-Aug-1996 #sequence-revision 16-Aug-1996 #text_change 17-Nov-2000
C:Accession: I40826; I40856; S51649
R:Hunter, J.E.C.; Brown, J.E.; Oyston, P.C.F.; Jakura, J.; Tibball, R.W.
Infect. Immun. 61, 3958-3965, 1993
A:Title: Molecular genetic analysis of the beta-toxin of Clostridium perfringens
A:Reference number: I40826; NUID:93366460; PMID:8359918
A:Accession: I40826
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-336 <RES>
A:Cross-references: GB:L13198; NID:G410018; PIDN:AAA23284.1; PID:G410019
R:Steinrdsdottir, V.; Frithrksdottir, V.; Gunnarsson, E.; Andresson, O.S.
FEMS Microbiol. Lett. 130, 273-278, 1995
A:Title: Expression and purification of Clostridium perfringens beta-toxin glutathione S
A:Reference number: I40856; NUID:93377614; PMID:7649450
A:Accession: I40856
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-170, 'K', 172-336 <RES>
A:Cross-references: EMBL:X83275; NID:G602993; PIDN:CAA58246.1; PID:G602994
C:Genetics:
A:Gene: cpb
C:Superfamily: leukocidin
C:Keywords: toxin

Query Match 19.7%; Score 293; DB 2; Length 336;
Best Local Similarity 25.2%; Pred. No. 2.2e-10;
Matches 80; Conservative 75; Mismatches 114; Indels 48; Gaps 12;
QY 13 DIGSNNTV---KTGD-----LVYD-----KENGHHKVPYFIDDKNHNK 51
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 29 DIGKTTTTRKNTSDGYTIITQNDKQIISYSDVSSSKNEDGFTASIDARFIDDKYSSEM 88
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 52 LLVIRTKGTIAG-----QYRVYSEEGANKSGLAWPSAFKVQ-LQLPDNEVAQISDYYP 103
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 89 TTLNLGTGFMSSKEDVIKKYNLH--DVTNSTAINFPVYISILNESINENVKIVDSIP 146
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 104 RNSIDTXEXMXTYXGXNXXYDX-XYTKXGXLGXNKSXGTYKXVQPFKFTILESPTD 162
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 147 KNTI--SQKTVSNMTGKIGSIEENKPKASIESEASESTIEYVQPDFSTIQTQDHS 204
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 163 KKGWKFVFNMMVQNGMPYDRDSWNPVYGNQLPMKTRNGSMKAADNPLDPNKAASSLLSS 222
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 205 SKASWDKFTETTR---GNYLNKSNPVYGNEMFYGRYTNVPATENLIIPDYQMSKLTG 261
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 223 GFSPDPAIVTMDRKASKQQTNDIVYERVDYQLHWTSTNMKGT-----NTKDKWT 275
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 262 GLNPNMSVVLTPN--GTEESIIKVKMERERCNYLNNWGANWYGVYSLRAFDTPN--V 317
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 276 DRSSERYKIDWEKEMT 292
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 318 DSHIFTKINLWTHKVT 334
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 15

JC5468

leukocidin chain lukM precursor - Staphylococcus aureus

C:Species: Staphylococcus aureus

C:Date: 07-Jul-1997 #sequence-revision 29-Aug-1997 #text_change 20-Jun-2000

C:Accession: JC5468; PC4338; S68460; S68542

R:Kaneko, J.; Muramoto, K.; Kamio, Y.

Biosci. Biotechnol. Biochem. 61, 541-544, 1997

A:Title: Gene of LukF-PV-like component of Panton-Valentine leukocidin in Staphylococ

A:Reference number: JC5468; MUID:97249691; PMID:9095557

A:Accession: JC5468

A:Molecule type: DNA

A:Residues: 1-308 <KAN>

A:Cross-references: DDBJ:D83951; NID:G1230553; PIDN:BAA12147.1; PID:G1230554

A:Experimental source: strain P83

A:Accession: PC4338

A:Molecule type: protein

A:Residues: 29-75; 101-130; 263-283 <KA2>

R:Choorit, W.; Kaneko, J.; Muramoto, K.; Kamio, Y.

FEBS Lett. 357, 260-264, 1995

A:Title: Existence of a new protein component with the same function as the LukF comp

A:Reference number: S68460; MUID:95137101; PMID:7835424

A:Accession: S68460

A:Molecule type: DNA

A:Residues: 1-2, 'N', 4-288, 'L', 290-299, 'S', 301, 'L', 303-304, 'R', 306-308 <CHO>

A:Cross-references: EMBL:DA2144; NID:G577648; PIDN:BAA07715.1; PID:G577649

A:Experimental source: strain P83 (ATCC 31890)

A:Accession: S68542

A:Molecule type: protein

A:Residues: 29-75; 101-130; 263-282 <CHW>

C:Genetics:

A:Gene: lukM

C:Function:

A:Description: causes cytotoxic changes in polymorphonuclear leukocytes in human and

C:Superfamily: leukocidin

C:Keywords: toxin

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-308/Product: leukocidin chain lukM #status experimental <MAT>

Query Match 17.7%; Score 262; DB 2; Length 308;

Best Local Similarity 24.4%; Pred. No. 1.6e-08;

Matches 71; Conservative 75; Mismatches 119; Indels 26; Gaps 12;

QY 9 TGTDDIGSNTTV--KTGDLVYTDKENGMMHKVVFYSFIDDKNNHKKLLVIRTKGTIAGQ-- 64
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |Db 30 TNAEDIGDDAEVIRKRTEDVSS--RKWGVTVQNVQDFVDRKKYKNDALIIKMQGFSINRST 87
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |QY 65 YRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDYYPNRSIDTXEXMXTYKXGXNXX 124
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |Db 88 FNDVKQNRANKR-MWPFQYNYGLTSKDQNTSLI-NYLPKNKIETVD--VGQTLGYNIGG 143
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |QY 125 TXDXTXKXGLXGXNXXSGXTXKXVQPDFKFTILESPTDKKVGWKFVFNMMV--NONNGPY 182
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |Db 144 KFQVSPISIGGSGSNYS--KSIKYSQKSYSEVEQSQSKTIKGVKANSFVIAGRHSAY 201
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |QY 183 DRDSWNPVYGNQLPMKTRNGSMKAADNPLDPNKAASSLLSSGFSPPDPAIVTMDRKASKQ 242
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |Db 202 D-----ELLFIRNTRGPNARDYFVDDNPLPLITSGFNPSFIATVSHE-KDSGDT 251
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |QY 243 TNIDVIYVRD-DYQLHWTSTNKGNTKDKWTDTRS-SERYKIDWEKEM 291
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |Db 252 SEFEITGRNMDVTYATYLPKGLGLYPERKNEFVNRNFVVKYEVNWKTYEI 302
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

Search completed: March 4, 2003, 10:45:10

Job time : 28.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:39:33 ; Search time 17 Seconds
(without alignments)
714.857 Million cell updates/sec

Title: CLAIM7
Perfect score: 1484
Sequence: 1 ADSIDINIKTGTDTIGSNTTV.....WDRSSERYKIDWEKEEMTN 293

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1484	100.0	319	1 HLA_STAAU	P09616 staphylococ
2	388	26.1	325	1 HLGE_STAAU	Q07226 staphylococ
3	374	25.2	323	1 LUKF_STAAU	P31715 staphylococ
4	253	17.0	309	1 HLGA_STAAU	P31714 staphylococ
5	248	16.7	315	1 HLGC_STAAU	Q07227 staphylococ
6	242	16.3	315	1 LUKS_STAAU	P31716 staphylococ
7	118	8.0	1584	1 KYK1_DICDI	P18160 dictyosteli
8	117.5	7.9	1787	1 CHD3_CABEL	Q22316 caenorhabdi
9	116.5	7.9	245	1 L43B_BOVIN	P29358 bos taurus
10	116	7.8	859	1 RPA2_METJA	Q58446 methanococc
11	115	7.7	461	1 EPIP_STAEP	P30199 staphylococ
12	114	7.7	464	1 MRJ4_APIME	Q17061 apis mellif
13	113.5	7.6	245	1 L43B_HUMAN	P31946 homo sapien
14	112	7.5	1183	1 CNA_STAAU	Q53654 staphylococ
15	111.5	7.5	1160	1 TFC3_YEAST	P34111 saccharomyc
16	111.5	7.5	1288	1 VACA_HELPJ	Q93455 helicobacte
17	110.5	7.4	245	1 L43B_MOUSE	Q9cqv8 mus musculu
18	109.5	7.4	798	1 PBPA_NEILA	Q85759 neisseria l
19	109	7.3	430	1 SYH_CHLPN	Q9z7p1 chlamydia p
20	109	7.3	1520	1 GLTB_BACSU	P39612 bacillus su
21	109	7.3	1986	1 WA_EWENI	Q03149 emericella
22	109	7.3	2334	1 WAPA_BACSU	Q07833 bacillus su
23	108.5	7.3	2477	1 FINC_MOUSE	P11276 mus musculu
24	108	7.3	1007	1 RGAL_YEAST	P39083 saccharomyc
25	107.5	7.2	329	1 PLC_BACTU	P08954 bacillus th
26	107.5	7.2	380	1 GLG2_YEAST	P47011 saccharomyc
27	107.5	7.2	429	1 DR48_YEAST	P18899 saccharomyc
28	107	7.2	430	1 YE09_YEAST	P40038 saccharomyc
29	107	7.2	1328	1 FINC_PLEWA	Q91289 pleurodeles
30	107	7.2	1562	1 YMB1_YEAST	Q04781 saccharomyc
31	106.5	7.2	493	1 GL03_YEAST	P38682 saccharomyc
32	106.5	7.2	843	1 YMS1_YEAST	Q05050 saccharomyc
33	106.5	7.2	2477	1 FINC_RAT	P04937 rattus norv

RESULT 1

ID	HLA_STAAU	STANDARD;	PRT;	319 AA.
AC	P09616;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Alpha-hemolysin precursor (Alpha-toxin) (Alpha-HL).			
GN	HLX OR HLA.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 27-44.			
RC	STRAIN=Wood 46;			
RX	MEDLINE=85053471; PubMed=6500704;			
RA	Gray G.S., Kehoe M.;			
RT	"Primary sequence of the alpha-toxin gene from Staphylococcus aureus			
RT	wood 46.";			
RL	Infect. Immun. 46:615-618(1984).			
RN	[2]			
RP	REVISIONS, SEQUENCE FROM N.A.			
RC	STRAIN=Wood 46;			
RA	Kedungrahn G.;			
RN	[3]			
RP	SEQUENCE OF 27-319 FROM N.A., AND PARTIAL SEQUENCE.			
RC	STRAIN=Wood 46;			
RX	MEDLINE=92268149; PubMed=1587866;			
RA	Walker B., Krishnaswamy M., Zorn L., Kasianowicz J., Bayley H.;			
RT	"Functional expression of the alpha-hemolysin of Staphylococcus			
RT	aureus in intact Escherichia coli and in cell lysates. Deletion of			
RT	five C-terminal amino acids selectively impairs hemolytic			
RT	activity.";			
RL	J. Biol. Chem. 267:10902-10909(1992).			
RN	[4]			
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).			
RC	STRAIN=Wood 46 / ATCC 10832;			
RX	MEDLINE=97102581; PubMed=8943190;			
RA	Song L., Hobaugh M.R., Shustak C., Cheley S., Bayley H., Gouaux J.E.;			
RT	"Structure of staphylococcal alpha-hemolysin, a heptameric			
RT	transmembrane pore.";			
RL	Science 274:1859-1866(1996).			
RN	[5]			
RP	MUTAGENESIS.			
RC	MEDLINE=93016135; PubMed=1400487;			
RA	Walker B., Krishnaswamy M., Zorn L., Bayley H.;			
RT	"Assembly of the oligomeric membrane pore formed by Staphylococcal			
RT	alpha-hemolysin examined by truncation mutagenesis.";			
RL	J. Biol. Chem. 267:21782-21786(1992).			
RN	[6]			
RP	MUTAGENESIS OF HISTIDINE RESIDUES.			
RC	MEDLINE=94222552; PubMed=8168947;			
RA	Menzies B.E., Kernodle D.S.;			
RT	"Site-directed mutagenesis of the alpha-toxin gene of Staphylococcus			
RT	aureus: role of histidines in toxin activity in vitro and in a murine			

34	106	7.1	517	1 Z215_HUMAN	Q9ul58 homo sapien
35	106	7.1	537	1 TEB6_STRPY	P18481 streptococc
36	106	7.1	576	1 CSG_METVO	Q50833 methanococc
37	105.5	7.1	245	1 L43B_RAT	P35213 rattus norv
38	105.5	7.1	928	1 HXA2_HAEIN	P45354 haemophilus
39	105	7.1	409	1 CAD1_YEAST	P24813 saccharomyc
40	104.5	7.0	638	1 GYRB_BACSU	P05652 bacillus su
41	104.5	7.0	655	1 CDGT_KLEPN	P08704 klebsiella
42	104.5	7.0	1145	1 POL_EIAVY	P03371 equine infe
43	104	7.0	311	1 PLC_STAAU	P45723 staphylococ
44	104	7.0	762	1 SLAP_ACEKI	P22558 acetogenium
45	104	7.0	928	1 NRPI_XENLA	P28824 xenopus lae

ALIGNMENTS


```
Db 200 SAYDOYLFAQDPTGP-AARDYFVDPNQLPPLIQSGFNPSFITTLSHERGKDKSE--FEI 256
QY 248 IYERVRDYLQHLWTSTNWKTKDKWTDERS-SERYKIDWEKEEM 291
Db 257 TYGRNMDATYAYTRHRLAVDRKHAFAKRNRTVTKYEVNWKTHEV 301

RESULT 5
HLGC_STAAU
ID HLGC_STAAU STANDARD; PRT; 315 AA.
AC Q07227;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-hemolysin component C precursor (H-gamma-1) (H-gamma-I).
GN HLGC.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93138806; PubMed=8423103;
RA Cooney J.C., Klenie Z., Foster T.J., O'Toole P.W.;
RT "The gamma-hemolysin locus of Staphylococcus aureus comprises three
RT linked genes, two of which are identical to the genes for the F and S
RT components of leukocidin."
RL Infect. Immun. 61:768-771(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89310396; PubMed=3075655;
RA Cooney J.C., Mulvey M., Arbutnot J., Foster T.;
RT "Molecular cloning and genetic analysis of the determinant for gamma-
RT lysin, a two-component toxin of Staphylococcus aureus."
RL J. Gen. Microbiol. 134:2179-2188(1988).
CC -1- FUNCTION: GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.
CC H-GAMMA-I IS ALSO THE F SUBUNIT OF LEUKOCIDIN.
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L01055; AAA26638.1; -
CC InterPro; IPR001776; Aerolysin.
CC InterPro; IPR001340; Hemolysn_pore.
CC InterPro; IPR003963; Staph_bicn_txn.
CC Pfam; PF01117; Aerolysin; 1.
CC PRINTS; PR01468; BICOMPFOXIN.
CC TIGRFAMS; TIGR01002; hlyII; 1.
CC Hemolysis; Toxin; Signal.
CC SIGNAL 1 29
CC CHAIN 30 315
CC SEQUENCE 315 AA; 35614 MW; 77359819736620BC CRC64;
CC -----
CC POTENTIAL.
CC GAMMA-HEMOLYSIN COMPONENT C.
CC FT CHAIN 30 315
CC SIGNAL 1 29
CC SEQUENCE 315 AA; 35614 MW; 77359819736620BC CRC64;
CC -----
Query Match 16.7%; Score 248; DB 1; Length 315;
Best Local Similarity 21.5%; Pred. No. 4.2e-08;
Matches 63; Conservative 85; Mismatches 117; Indels 28; Gaps 11;
QY 11 TTDIGSTNTVTKTGDV--TYDKEN---GMHKVFYFSDIDKHNKLLVTRTKGTAGQY 65
| | | | : : : | : : | : : | : : | : : | : : | : : | : : | : :
| | | : : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 33 TEDIGKSGSDI---EIKRTEDKTSNKGVTQNIQDFVCKDKYNKDAILKMQGFISRT 89
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
QY 66 RVYSEEGANK-SGLAWPSAFKVLQPLDNEVAOISDYYPENSIDTXEMXTYTXGXNXX 124
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 90 TYNYKNTNHVKARWPFQNIIGLTKNDKYSLI-NYLPKNKIESTN--VSQTLGNIGG 146
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
QY 125 TXDXTKXGXGLXGNXSXGTXKXKXVQPDFKLTILESPTDKKVGKVFNNMNQWGPYDR 184
```

```
Db 147 NFOSAPSLGNGSFNYS--KSISYTOQNYVSEVEQNSKSVLWGVKANSFATES----- 198
QY 185 DSNVPYGNQLEFKTRNGSKMKAADNFDPNKASSLLSSGSPDFAITVITMDRKASKOOTN 244
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 199 -GOKSAFSDSLFVGYKPHSKDPRDYFVPDSELPPLVOSGFNPFIATVSHE-KGSDTSE 256
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
QY 245 IDVIYERVRD-----DYQLHWTSTNWKTKDKWTDERS-SERYKIDWEKEEM 291
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 257 FEITYGRNMDVTTHAIKRSTHYGNSYLDGHRVHNAFVNRNRYTVKYEVNWKTHEI 309
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : :

RESULT 6
LUKS_STAAU
ID LUKS_STAAU STANDARD; PRT; 315 AA.
AC P31716;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leukocidin S subunit precursor.
GN LUKS.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-79.
RX STRAIN=NO.4;
RX MEDLINE=92068180; PubMed=1958181;
RA Rahman A., Izaki K., Kato I., Kamio Y.;
RT "Nucleotide sequence of leukocidin S-component gene (luks) from
RT methicillin resistant Staphylococcus aureus."
RL Biochem. Biophys. Res. Commun. 181:138-144(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=MRSA NO. 4;
RX MEDLINE=93372503; PubMed=7763998;
RA Rahman A., Izaki K., Kamio Y.;
RT "Gamma-hemolysin genes in the same family with lukF and luks genes in
RT methicillin resistant Staphylococcus aureus."
RL Biosci. Biotechnol. Biochem. 57:1234-1236(1993).
CC -1- FUNCTION: LEUKOCIDIN CAUSES CYTOTOXIC CHANGES IN POLYMORPHONUCLEAR
CC LEUKOCYTES.
CC -1- SUBUNIT: LEUKOCIDIN CONSISTS OF TWO PROTEIN COMPONENTS: F AND S.
CC -1- SIMILARITY: TO THE H-GAMMA II SUBUNIT OF GAMMA-HEMOLYSIN.
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M81346; AAA26654.1; -
CC EMBL; S65052; AAC60445.1; -
CC PIR; JN0626; JN0626.
CC InterPro; IPR001776; Aerolysin.
CC InterPro; IPR001340; Hemolysn_pore.
CC InterPro; IPR003963; Staph_bicn_txn.
CC Pfam; PF01117; Aerolysin; 1.
CC PRINTS; PR01468; BICOMPFOXIN.
CC TIGRFAMS; TIGR01002; hlyII; 1.
CC Hemolysis; Toxin; Signal.
CC SIGNAL 1 29
CC CHAIN 30 315
CC SEQUENCE 315 AA; 35557 MW; 6F480166DF13AA1E CRC64;
CC -----
Query Match 16.3%; Score 242; DB 1; Length 315;
Best Local Similarity 21.2%; Pred. No. 1e-07;
Matches 62; Conservative 85; Mismatches 118; Indels 28; Gaps 11;
QY 11 TTDIGSTNTVTKTGDV--TYDKEN---GMHKVFYFSDIDKHNKLLVTRTKGTAGQY 65
```

```
Db 33 TEDIGKSDI---EIKRTEDTSNKGWYTONIQDFVDRKYNKDALLIKMQGFSSRT 89
QY 66 RYSEEGANK-SGLAWPSAFKVLQLPDNEVAQISDYPRNSIDIXEXMXTYXGNXK 124
Db 90 TTYNYKKTTHVAMRWPFOYNGITGLTNDYVSLI-NYLPKNKIESTN--VSQTLGYNIGG 146
QY 125 TXDXTKXGKXGNXSGXTKXKXQVDPDKFTILESPDKKKVGKVIFFNNVNNQNWGPYDR 184
Db 147 NQOSAPSLGSGSNFY--KSISYQONVYVEQONSXVLWGKANSFATES----- 198
QY 185 DSNPNVYGNQLPMKTRNGSKMAADNPLDPNKASSLLSGFSPDFATVITMDRKASKQQTIN 244
Db 199 -GOKSAFDSDLFVGKPHSKDPRDYFVPDSELPPLVQSGFNPFSFIATVSHE-KGSDTSE 256
QY 245 IDVIVYERVD-----DYQLHWTSTNKKGTNTKDKWTDTRS-SERYKIDWEKEEM 291
Db 257 FEITYGRNMDVTHAKRSTRHYGNSYLDGHRVHNAFVNRNRYTYKYEYVNWKTHEI 309

RESULT 7
ID KYKL_DICDI STANDARD; PRT; 1584 AA.
AC P18160;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-
protein kinase 1).
GN PYKA OR SPLA OR DPYKL.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Metazoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JH10;
RX MEDLINE=97053827; PubMed=8998241;
RA Nuckolls G.H., Osheroov N., Loomis W.F., Spudich J.A.;
RT "The Dictyostelium dual-specificity kinase spla is essential for
spore differentiation.";
RL Development 122:3295-3305(1996).
RN [2]
RP SEQUENCE OF 1248-1584 FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
Dictyostelium discoideum.";
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -!- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
tyrosine phosphate.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK
DURING THE MOUND STAGE OF MORPHOGENESIS.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; U32174; AAB41125.1; -
DR EMBL; M33785; AAA33202.1; -
DR PIR; A35670; A35670.
DR Dictydb; DD03010; pyka.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR003878; SPRY_domain.
DR InterPro; IPR003877; SPRY_receptor.
DR InterPro; IPR004040; STY_pkinase.
```

```
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR Pfam; PF00622; SPRY; 3.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00449; SPRY; 3.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00105; SAM_DOMAIN; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT DOMAIN 908 972 SAM.
FT DOMAIN 403 420 POLY-ASN.
FT DOMAIN 428 435 POLY-THR.
FT DOMAIN 449 480 POLY-ASN.
FT DOMAIN 483 491 POLY-ASN.
FT DOMAIN 494 508 POLY-ASN.
FT DOMAIN 512 532 POLY-ASN.
FT DOMAIN 596 600 POLY-ASN.
FT DOMAIN 808 811 POLY-PHE.
FT DOMAIN 1026 1029 POLY-SER.
FT DOMAIN 1195 1210 POLY-ASN.
FT DOMAIN 1215 1220 POLY-GLN.
FT DOMAIN 1224 1233 POLY-GLN.
FT DOMAIN 1266 1274 POLY-PRO.
FT DOMAIN 1289 1561 PROTEIN_KINASE.
FT NP_BIND 1295 1303 ATP (BY SIMILARITY).
FT BINDING 1316 1316 ATP (BY SIMILARITY).
FT ACT_SITE 1417 1417 D -> R (IN REF. 2).
FT CONFLICT 1248 1248 D -> R (IN REF. 2).
FT CONFLICT 1435 1435 V -> L (IN REF. 2).
SQ SEQUENCE 1584 AA; 174304 MW; 5D1589458D8E01E3 CRC64;

Query Match 8.0%; Score 118; DB 1; Length 1584;
Best Local Similarity 20.7%; Pred. No. 60;
Matches 36; Conservative 48; Mismatches 50; Indels 40; Gaps 6;

QY 93 NEVAQI-----SDYVPRNSIDTXEXMXTYXGNXNXTXDTKXKXGLXGNXSGXTXK 147
Db 967 NRMQIWNKSPDYPKTAIDSSDKIRWPASGGSGG---INISGVVIGSSGSDGGIT 1023
QY 148 XVQPDFKTI--LESPTDKKVGKVIFFNNVNNQNWGPYDRDSNPNVYG----NOLFMTNRN 201
Db 1024 EISSSSKNIRPVKSYTQKEI-----EDNRNSTISGGEKKNKYI----- 1063
QY 202 GSKAADNPLDPNKASSLLSGFSPDPATVITMDRKASKQQTINIDVIVYERVD 255
Db 1064 -----DNQMPHQIGSDGLLPDFGQPPDEKNSKSTLSNEQIRYLQQRKD 1111

RESULT 8
CHD3_CAEEL
ID CHD3_CAEEL STANDARD; PRT; 1787 AA.
AC Q22516; Q18794;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chromodomain helicase-DNA-binding protein 3 homolog (CHD-3).
GN CHD-3 OR TL468.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20530482; PubMed=11076750;
RA von Zelewsky T., Palladino F., Brunschwig K., Tobler H., Hajnal A.,
RA Mueller F.;
RT "The C. elegans Mi-2 chromatin-remodelling proteins function in vulval
cell fate determination.";
RL Development 127:5277-5284(2000).
```



```

FT CHAIN 76 546 MJA RPOA2 INTEIN (SPACER PROTEIN)
FT CHAIN 547 859 RPOA2, 2ND PART (POTENTIAL)
FT CHAIN 547 859 RPOA2, 2ND PART (POTENTIAL)
SQ SEQUENCE 859 AA; 97086 MW; A05799E007899015 CRC64;

Query Match
Best Local Similarity 7.8%; Score 116; DB 1; Length 859;
Matches 58; Conservative 53; Mismatches 99; Indels 36; Gaps 13;

QY 26 VYDKENGHHKVFYFIDDKHKKLLVIRKQ---TIAGQYRVYSEEGANK-----S 76
DB 127 LSLDQDEKQVHWRKLIISCIHKK-HNGKLIKIKTKSGREITATPYHSFVIRKDKNIIPVKGS 185

QY 77 GLAWPSAFVQLQLPDN--EVAQISDYPRNSI--DTXEXMXTXTXGXNKNXTXDXTXKX 132
DB 186 ELKIGRIDPVVHHIPANCVAEINISDVSGNVVDNINNKIAPKNGSIPNNIKLIDYDF 245

QY 133 GXLGXNKSXGTXKXVQDPFKTILESPTDKKVGKVV--IFNNMNVGNWGPYDRDSWNP-- 189
DB 246 GYFIGIYLAEGSVTKY---FVSI--SNVDELILNKIRAFADKLGILNYGEYDNNNGFAES 299

QY 190 ----VYGNOL--FMK---TRNSGMAADNFDLPNK--ASSLLSGSPSPFAVITMDRKA 238
DB 300 HDIRYSSVTLAEFLSNFTSGTSNTKIAEFVFGANKFEVRLIRGYPDGD--GNVNA DRKV 357

QY 239 SKOOTN 244
DB 358 IRVTSN 363

RESULT 11
EPIP_STAEP STANDARD; PRT; 461 AA.
AC P30199.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermin leader peptide processing serine protease EPIP precursor
DE (EC 3.4.21.-)
GN EPIP.
OS Staphylococcus epidermidis.
OG Plasmid pTu 32.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-TU 3298 / DSM 3095;
RX MEDLINE=92155237; PubMed=1740156;
RA Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
RA Goetz F., Entian K.-D.;
RT "Analysis of genes involved in the biosynthesis of lantibiotic
RT epidermin."
RL Eur. J. Biochem. 204:57-68(1992).
CC -!- FUNCTION: PROTEASE WHICH CLEAVES THE MATURED LANTIBIOTIC FROM THE
CC MODIFIED PREPEPTIDE (PROBABLY).
CC -!- PATHWAY: Epidermin biosynthesis; last step.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X62386; CAA44257.1;
DB PIR; S23420; S23420.
DB HSSP; P00782; 2SBT.
DB MEROPS; S08.060; -.
DB InterPro; IPR000209; Peptidase_S8.
DB Pfam; PF00082; Peptidase_S8; 1.
DB PRINTS; PR00723; SUBTILISIN.

```

```

DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KW Hydrolase; Serine protease; Signal; Zymogen; Plasmid.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 ?
FT CHAIN ? 461
FT EPIDERMIN LEADER PEPTIDE PROCESSING
FT SERINE PROTEASE EPIP.
FT ACT_SITE 149 149 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 194 194 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 402 402 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 461 AA; 51814 MW; F2770F4F8436D906 CRC64;

Query Match
Best Local Similarity 7.7%; Score 115; DB 1; Length 461;
Matches 63; Conservative 57; Mismatches 118; Indels 52; Gaps 13;

QY 4 DINIKGTTDIGS-----NTTVKTKGDLVTYDKENGHHKVKVFYFIDDKHNNKLLV--IR 56
DB 149 DSGVNSSHTDLKSINKIVNEVPKNGFRGSENDESGNK-----NFEEDKLNHGTLVAGQIG 203

QY 57 TKGTTAG-----QYRVYSEEGANKSLAWPSAFKVLQLPDNEVAQIS-----DYYP 104
DB 204 ANGNLKGVNPGVEMNMYRVFGSKSEMLWVSGIIDAANDNDVINVSGLNYLIKDNQNK 263

QY 105 NSIDTQEXMXTXTGXNKNXTXKXCLXGXNKSXGTXKXVQ--DFKTILESPDK 163
DB 264 KKLDRDEKVDYDALOKAINTAQ---KGSIVVAAVGNCGDGINVKVKEINKRNLSKTSK 320

QY 164 KVGWKVIFNNMVNQN---WGYPYDRDSWNPV---YGNQLF-MKTRNGSMKAADNFDLPNK 215
DB 321 KV-----YDSPANLNVMVTGSDIDNDYISEFSNYGNNEIDLMTIGGSYK----LLDKYG 371

QY 216 ASSLLSGSPSPFAVITMDRASKQQTWIDYIYERV-----DDYQL 258
DB 372 KQAWLEKGYMQKQSVLSTSSNGRIYQSGTSLAAPKVGALALEIDKYQL 421

RESULT 12
MRJ4_APIME STANDARD; PRT; 464 AA.
AC Q17061;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major royal jelly protein 4 precursor (MRJP-4) (Bee-milk protein)
DE (Royal jelly protein RJP57-2).
GN MRJP4.
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Apoidea; Apidae; Apis.
OX NCBI_TaxID=7460;
[1]
RN TISSUE=Head;
RP SEQUENCE FROM N.A.
RA Klaudiny J., Hanes J., Kulifajova J., Albert S., Slimuth J.;
RT "Molecular cloning of two cDNAs from the head of the nurse honey bee
RT (Apis mellifera L.) for coding related proteins of royal jelly.";
RL J. Apicultural Res. 33:105-111(1994).
[2]
RP CHARACTERIZATION.
RC TISSUE=Head;
RX MEDLINE=99007754; PubMed=9791542;
RA Hanes J., Judova J., Slimuth J.;
RT "A family of major royal jelly proteins of the honeybee Apis mellifera
RL L.";
RL Cell. Mol. Life Sci. 54:1020-1030(1998).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN HONEYBEE NUTRITION. IT IS
CC FOUND IN THE ROYAL JELLY WHICH IS THE FOOD OF THE QUEEN HONEY BEE
CC LARVA. THE ROYAL JELLY DETERMINES THE DEVELOPMENT OF THE YOUNG
CC LARVAE AND IS RESPONSIBLE FOR THE HIGH REPRODUCTIVE ABILITY OF THE

```



```

DR EMBL; AL008725; CAA15497.1; -
DR EMBL; BC001359; AAH01359.1; -
DR HSP; P29312; 1A38.
DR Genew; HGNC:12849; YWHAB.
DR MIM; 601289; -
DR InterPro; IPR000308; 14-3-3-3.
DR Pfam; PF00244; 14-3-3; 1.
DR PRINTS; PR00305; 1433ZETA.
DR ProDom; PD000600; 14-3-3; 1.
DR SMART; SM00101; 14-3-3; 1.
DR PROSITE; PS00796; 1433_1; 1.
DR PROSITE; PS00797; 1433_2; 1.
KW Brain; Neurone; Phosphorylation; Acetylation; Multigene family;
FT Alternative initiation.
FT INIT_MET 0
FT CHAIN 1 245
FT CHAIN 2 245
FT INIT_MET 2 2
FT MOD_RES 1 1
FT MOD_RES 2 2
FT MOD_RES 185
FT MOD_RES 245 AA; 27951 MW; OBAS9BF97595485 CRC64;
SQ SEQUENCE 245 AA; 27951 MW; OBAS9BF97595485 CRC64;

Query Match 7.6%; Score 113.5; DB 1; Length 245;
Best Local Similarity 19.1%; Pred. No. 6.1;
Matches 50; Conservative 52; Mismatches 87; Indels 73; Gaps 9;

QY 28 YKENGKMKVYFSDDKHNKLLVIRTKGTIAQ---YRVYS-----BEGANKSGL 78
DB 20 YDMAAMKAVTEQGHLSNEERLLSVAYKNVGARRSWRVSSIEOKTERNEKKQOM 79

QY 79 AMPFAFKVOLQLPD--NEVAQISDYI-----PRNSI-----DTXEXMTXTXGXNX 122
DB 80 GREYREKIEAEQLQICNDVLELLDKYLIPNATQPKSVKYLKMGDYFYLSEVASGDNK 139

QY 123 NVTXDYTKXGKXGXNKSAGYTXKXVQDPFTILSPDVKVGVKVFIPNNMVNONWGPY 182
DB 140 QTTVSNQQA-----YQEAFFISKEMQPTPIRLGLALNFSVFYEILNS----- 185

QY 183 DRDSWNPVYGQLEFMKTRNGSKAADNFDLPNKASSLLSSGFSPOFATVITMDRKASKQ 242
DB 186 -----PEKACSLAKTAFDEATAELDTLNEESYKDS 215

QY 243 TNIDVIYERVDYQLHWTSTN 264
DB 216 T---LIMQLLRNLTL-WTSEN 233

RESULT 14
CNA_STAAU STANDARD; PRT; 1183 AA.
AC Q53654;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen adhesin precursor.
GN CNA.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FDA 574;
RX MEDLINE=92165839; PubMed=1311320;
RA Patti J.M., Jonsson H., Guss B., Switalski L.M., Wiberg K.,
RA Lindberg M., Hoeek M.;
RT "Molecular characterization and expression of a gene encoding a
RT Staphylococcus aureus collagen adhesin.";
RL J. Biol. Chem. 267:4766-4772(1992).
RN [2]
RP ERRATUM.
RA Patti J.M., Jonsson H., Guss B., Switalski L.M., Wiberg K.,

```

```

RA Lindberg M., Hoeek M.;
RA J. Biol. Chem. 267:11672-11672(1994).
RN [3]
RP COLLAGEN-BINDING DOMAIN.
RC STRAIN=FDA 574;
RX MEDLINE=94032261; PubMed=8218209;
RA Patti J.M., Boles J.O., Hoeek M.;
RT "Identification and biochemical characterization of the ligand
RT binding domain of the collagen adhesin from Staphylococcus aureus.";
RL Biochemistry 32:11428-11435(1993).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 169-318.
RX MEDLINE=97475225; PubMed=9334749;
RA Symersky J., Patti J.M., Carson M., House-Pompeo K., Teale M.,
RA Moore D., Jin L., Schneider A., DeLucas L.J., Hoeek M.,
RA Narayana S.V.L.;
RT "Structure of the collagen-binding domain from a Staphylococcus
RT aureus adhesin.";
RL Nat. Struct. Biol. 4:833-838(1997).
CC -!- FUNCTION: MEDIATES ATTACHMENT OF STAPHYLOCOCCAL CELLS TO
CC COLLAGEN-CONTAINING SUBSTRATA.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL; M81736; AAA20874.1; -
DB PDB; 1AMX; 24-JUN-98.
DR InterPro; IPR001899; Gram_pos_anchor.
DR TIGRfams; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; FALSE_NEG.
KW Cell wall; Peptidoglycan-anchor; Repeat; Signal; 3d-structure.
FT SIGNAL 1 29
FT CHAIN 30 1154
FT PROPEP 1155 1183
FT DOMAIN 151 318
FT DOMAIN 533 1093
FT DOMAIN 1093 1157
FT REPEAT 533 719
FT REPEAT 720 906
FT REPEAT 907 1093
FT SITE 1151 1155
FT MOD_RES 1154 1154
FT SEQUENCE 1183 AA; 133066 MW; B6A1CC072E57D76 CRC64;

Query Match 7.5%; Score 112; DB 1; Length 1183;
Best Local Similarity 18.9%; Pred. No. 90;
Matches 60; Conservative 62; Mismatches 118; Indels 78; Gaps 12;

QY 3 SDINIKTGTDIGSTNTVTKTGLVTVYDKENGHMKVYFSDDKHNKLLVIRTKGTIA 62
DB 39 TDLTVSPSKIEDGGKTVK---WTFDDKNG-----KIQGDMIKV----- 75

QY 63 GOYRVYSEEGANKSGLAWPSAFKVLQI-----LPDNEVAQISDYI 102
DB 76 -----AMPTSGTVKIEGYSKTVPLTVKGEQGVQAVITPDGATITFNDKV 119

QY 103 PRNSIDTXXEMXTYTXGXNKNXTXDTXKXGLXGNXSGTXKXVQDP-----FKTI 156
DB 120 EKLX-DVSGFAEFVQGRNLTQNTSDDKVATITSGNKTNTVHKSEAGTSVFFYKGT 178

QY 157 LESPTD-KKVGKVIFFNNMVNONWGPYDRDSWNPVY-GNOLFMTKTRNGSKAA-DNFLDP 213
DB 179 DMLPETHVRWFLINN--EASVYSKDITIKDQIQGGQQLDLSTLININVTGTHSNYSG 236

QY 214 NKASSLLSSGFSPOFATVITMDRKASKQQTNDIYERVDYQLHWTSTNMGKTNTKDK 273

```

```
Db 237 QSAITDEKAFP---GSKITVD-----NTKNTIDVTIPOGYSYNSFSINYTKITNEQOK 289
Qy 274 WTDSSERYKIDWEKEEM 291
Db 290 EFVNSQAWQEHGKEEV 307

RESULT 15
TFC3_YEAST
ID TFC3_YEAST STANDARD; PRT; 1160 AA.
AC P34111;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor tau 138 kDa subunit (TFIIIC 138 kDa subunit).
GN TFC3 OR TSV115 OR YAL001C OR FUN24.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=S288c;
RX MEDLINE=93066269; PubMed=1279682;
RA Lefebvre O., Carles C., Conesa C., Swanson R.N., Bouet F., Riva M.,
RA Sentenac A.;
RT "TFC3: gene encoding the B-block binding subunit of the yeast
RT transcription factor IIIC.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10512-10516(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=95028152; PubMed=7941740;
RA Clark M.W., Keng T., Storms R.K., Zhong W., Fortin N., Zeng B.,
RA Delaney S., Ouellette B.F.F., Barton A.B., Kaback D.B., Bussey H.;
RT "Sequencing of chromosome I of Saccharomyces cerevisiae: analysis of
RT the 42 kbp SPO7-CEN1-CDC15 region.";
RL Yeast 10:535-541(1994).
CC -|- FUNCTION: TFIIIC MEDIATES TRNA AND 5S RNA GENE ACTIVATION BY
CC BINDING TO INTRAGENIC PROMOTER ELEMENTS. IT ASSEMBLES THE
CC INITIATION COMPLEX TFIIIB-TFIIIC-TDNA. TFC3 IS ESSENTIAL FOR
CC CELL VIABILITY.
CC -|- SUBUNIT: COMPOSED OF AT LEAST FOUR DIFFERENT SUBUNITS.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- SIMILARITY: SOME, TO S.POMBE SPBC336.07.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M98261; AAA34378.1; -.
CC EMBL; L22015; AAC04956.1; -.
CC PIR; A46423; A46423.
CC PIR; S40892; S40892.
CC PIR; S43444; S43444.
CC TRANSPAC; T03679; -.
CC SGD; S0000001; TFC3.
CC Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 1 68 HMG BOX.
FT DNA_BIND 1037 1110 HMG BOX.
SQ SEQUENCE 1160 AA; 132108 MW; 3ACB7893ED536581 CRC64;

Query Match 7.5%; Score 111.5; DB 1; Length 1160;
Best Local Similarity 16.1%; Pred. No. 93;
Matches 66; Conservative 83; Mismatches 114; Indels 147; Gaps 15;

Qy 1 ADSDINKTGTDTGSGNTVTGKGLVTDYDKEN--GMHKVYFYSFIDD----- 45
Db 418 AEDEISVPKGFDELGSRT----DLKTLNEDNFVALNTRVFTTDSGDQDIFFWHGELKI 473
```

```
Qy 46 -----KNHKKLLVIRTKGTIAGQYRVYSEEGANKSLWAPSAFKVOLQLPDNEVA 96
Db 474 PPNSKKTTPNKNRRQVKNSTNASVAG-----NISPKRIKLEQHVSTAQEP 520
Qy 97 QISDYYPNRSIDT-XEXMXTXTXGXNXXNTXDXTXKXGLXGXNKSXGXT----- 145
Db 521 KSAEDSPSSNGGTVVKGKVVNFGFSARSLRSLQRRAILKVMNTIGGVAYLREQFYEV 580
Qy 146 -----KXVQDPDKTILESP-----TDKKVGVKVIENNMVN-----QNKGPYDRD 185
Db 581 SKYMGSTTTTLDKKTVRGDDVLDLWVESEKLGARTEPVSGRKIIIFLPTVGEDAIQRILEKD 640
Qy 186 S-----WNPVYGNOLEM-----KTRNGSMKAADNFDLPNKA 216
Db 641 SKKATFTDVIHDTIYFFDQTEKRNFRHGKKSVERIRKFNQRQNAKIKASDDAISKKST 700
Qy 217 SSLSSGSPDFATVITMDRKASKQOQTNIDVIYERVRDDYOLHWTST----- 263
Db 701 SVNVDG-----KIKRRDKKVSAGRTT--VVVENTKEDKTVYHAGTKDGVQALIRAVV 752
Qy 264 -----NWKGTNTKDKWTDR-----SSERYKID-WEK 288
Db 753 TKSINKIEMWDKITKLPNNSLDNLKKKKTARRVYRMGHSGHWRAYVDKWK 802
```

Search completed: March 4, 2003, 10:43:08
Job time : 20 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:40:04 ; Search time 28.5 seconds
(without alignments)
2118.309 Million cell updates/sec

Title: CLAIM7

Perfect score: 1484

Sequence: 1 ADSIDINIKGTDTIGSNVTY.....WDRSSRYKIDWEKEETN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1480	99.7	319	16	Q99uu6 staphylococ
2	437.5	29.5	412	2	O05387 bacillus ce
3	401	27.0	325	2	O50604 staphylococ
4	397	26.8	325	9	O80067 staphylococ
5	396	26.7	327	2	O54082 staphylococ
6	395.5	26.7	327	2	Q93uu8 staphylococ
7	394.5	26.6	327	16	Q99T54 staphylococ
8	394.5	26.6	332	16	Q931n4 staphylococ
9	394	26.5	325	2	O53747 staphylococ
10	393	26.5	325	2	O57227 staphylococ
11	385	25.9	325	16	Q931f3 staphylococ
12	385	25.9	326	16	Q99sn8 staphylococ
13	384.5	25.9	326	2	O54327 staphylococ
14	374	25.2	322	2	O53704 staphylococ
15	374	25.2	322	9	Q9mbn2 staphylococ
16	351	23.7	336	2	Q9eval bacillus ce

17	314	21.2	265	2	Q937V1	Q937v1 bacillus ce
18	313	21.1	265	2	Q937V0	Q937v0 bacillus ce
19	308.5	20.8	260	2	Q937V2	Q937v2 bacillus ce
20	293	19.7	336	2	O46308	O46308 clostridium
21	293	19.7	336	2	O46181	O46181 clostridium
22	289	19.5	309	2	O91403	O91403 clostridium
23	265	17.9	308	2	O53731	O53731 staphylococ
24	262	17.7	308	2	O53703	O53703 staphylococ
25	262	17.7	308	9	Q9MEN3	Q9mbn3 staphylococ
26	256.5	17.3	351	16	O931i5	O931i5 staphylococ
27	254.5	17.1	351	16	Q99SN7	Q99sn7 staphylococ
28	251	16.9	315	2	O53701	O53701 staphylococ
29	249	16.8	286	2	O9AFA9	O9afa9 staphylococ
30	248	16.7	286	2	O9AFB0	O9afb0 staphylococ
31	248	16.7	315	2	O53691	O53691 staphylococ
32	248	16.7	315	16	O99RL1	O99rl1 staphylococ
33	247.5	16.7	311	2	Q93UU9	Q93uu9 staphylococ
34	247.5	16.7	311	16	Q99T53	Q99t53 staphylococ
35	239	16.1	314	2	O54081	O54081 staphylococ
36	237	16.0	312	2	O50603	O50603 staphylococ
37	237	16.0	312	9	O80066	O80066 staphylococ
38	231	15.6	312	2	O53746	O53746 staphylococ
39	203.5	13.7	310	2	O54326	O54326 staphylococ
40	147	9.9	28	2	P97144	P97144 staphylococ
41	129	8.7	1272	16	Q98PR8	Q98pr8 mycoplasma
42	127.5	8.6	1348	5	O8T234	O8t234 dictyostelli
43	125.5	8.5	290	5	O94482	O94482 dictyostelli
44	124	8.4	410	9	O92XE1	O92xel bacterioph
45	119	8.0	447	2	Q9RPL5	Q9rpl5 streptococ

ALIGNMENTS

RESULT 1

Q99UU6	PRELIMINARY;	PRT;	319 AA.
AC	Q99UU6		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Alpha-hemolysin precursor.		
GN	SAV1163 OR SA1007.		
OS	Staphylococcus aureus (strain Mu50 / ATCC 700699), and		
OS	Staphylococcus aureus (strain N315).		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Staphylococcus.		
OX	NCBI_TaxID=158878, 158879;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPECIES=Staphylococcus aureus (strain Mu50), and Staphylococcus aureus (strain N315);		
RX	MEDLINE=21311952; PubMed=11418146;		
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,		
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,		
RA	Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,		
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,		
RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,		
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,		
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;		
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus aureus."		
RL	Lancet 357:1225-1240(2001).		
DR	EMBL; AF003361; BAB57325.1; -		
DR	EMBL; AF003312; BAB42258.1; -		
DR	HSSP; P09616; 7AHL.		
DR	InterPro; IPR001776; Aerolysin.		
DR	InterPro; IPR001340; Hemolysin_pore.		
DR	InterPro; IPR003963; Staph_bicn_txn.		
DR	Pfam; PF01117; Aerolysin, I.		
DR	PRINTS; PR01468; BICOMPFOXIN.		
DR	TIGRFAMS; TIGR01002; hlyII; 1.		
DR	PROSITE; PS00274; AEROLYSIN; 1.		
KW	Complete proteome.		

```
SQ SEQUENCE 319 AA; 35975 MW; 9AB4C82FB2CD9698 CRC64;
Query Match 99.7%; Score 1480; DB 16; Length 319;
Best Local Similarity 92.8%; Pred. No. 3.8e-80;
Matches 272; Conservative 21; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADSINIKTGTGIDGNTTWTGDLVTDKENGHMKKVFYSFIDDKHNNKLLVIRTKGT 60
Db 27 ADSINIKTGTGIDGNTTWTGDLVTDKENGHMKKVFYSFIDDKHNNKLLVIRTKGT 86
QY 61 IAGQRYVSEEGANKSLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXXMXXTXGX 120
Db 87 IAGQRYVSEEGANKSLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXXMXXTXGX 146
QY 121 NXNXTYDXTXKXGALGXNXXSXTXKXVQDPDKTILESPTDKKVGKVFNNMWNQWG 180
Db 147 NGNVTGDTGKIGGLIGANYSIGHTLKYVQDPDKTILESPTDKKVGKVFNNMWNQWG 206
QY 181 PYDRDSNPNVYGNQLEFMKTRNGSKAADNFDLPNKASSLLSGFSDFATVITMDRKSK 240
Db 207 PYDRDSNPNVYGNQLEFMKTRNGSKAADNFDLPNKASSLLSGFSDFATVITMDRKSK 266
QY 241 QQTNDIVIVRVDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEEMTN 293
Db 267 QQTNDIVIVRVDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEEMTN 319

RESULT 2
O05387 PRELIMINARY; PRT; 412 AA.
AC O05387;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hemolysin II.
GN HLX-II.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VKM-B771;
RX MEDLINE=94111714; PubMed=8283976;
RA Sinev M.A., Budarina Z.I., Gavrilenko I.V., Tomashevskii A.I.,
Kuzmin N.P.;
RT "Evidence of the existence of hemolysin II from Bacillus cereus:
cloning the genetic determinant of hemolysin II."
RL Mol. Biol. (Mosk.) 27:1218-1229(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VKM-B771;
RX MEDLINE=20016563;
RA Baida G., Budarina Z.I., Kuzmin N.P., Solonin A.S.;
RT "Complete nucleotide sequence and molecular characterization of
hemolysin II gene from Bacillus cereus."
RL FEMS Microbiol. Lett. 180:7-14(1999).
DR EMBL; U94743; AAB51536.1; -
DR HSSP; P09616; 7AHL.
DR InterPro; IPR001776; Aerolysin.
DR Pfam; PF01117; Aerolysin_pore.
DR TIGRFAMs; TIGR01002; hlyII; 1.
SQ SEQUENCE 412 AA; 45664 MW; D7E8313151C85A8F CRC64;

Query Match 29.5%; Score 437.5; DB 2; Length 412;
Best Local Similarity 31.9%; Pred. No. 7.6e-18;
Matches 88; Conservative 71; Mismatches 96; Indels 21; Gaps 7;

QY 27 TYDENGHMKKVFYSFIDDKHNNKLLVIRTKGT-IAGQRYVSEEGANKSLAWPSAFK 85
Db 53 TYDMKQNIKSIKVSFIEDPYADKKIAIVTDDGSNIDAKYTI--NGGYNAGLWPSAYH 110
```

```
QY 86 VQLQLPDNEVAQISDYPRNSIDTXXMXXTXGXNXXTXDXTXXKXGL-XGXNXSGX 144
Db 111 TEAEITSGDSAQFHKAAPVN-----TMTSAKYTSEVGYTLGSGVKGVNDKGPNDASI 164
QY 145 TXXV-----OPDKTILESPTDKKVGKVFNNMWNQNGPYDRDSNPNVYGNQLF 196
Db 165 TGSFAWKESYSDQVDYKTVLETHDKLKNKVGFSQSFNPEWGIYNRDSFNTFYGNLF 224
QY 197 MKYRNGSKMAADNFDLPNKASSLLSGFSDFATVITMDRKSKQQTNDIVIVRVDY 256
Db 225 MKRSYN-EGTNFVSKDTPVALTGFSNVAIVATDK--TESTSCLKITNRRISDQY 281
QY 257 QLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEEMT 292
Db 282 NIEWSSKMGWGTNNKDTYNEFFTNKYKLDWKNHQVT 317

RESULT 3
O50604 PRELIMINARY; PRT; 325 AA.
AC O50604;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE LukF-PV.
GN LukF-PV.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC49775;
RX MEDLINE=98067870; PubMed=9404084;
RA Kaneko J., Muramoto K., Kamio Y.;
RT "Gene of LukF-PV-like component of Panton-Valentine leukocidin in
Staphylococcus aureus P83 is linked with lukM."
RL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).
DR EMBL; AB006796; BAA24008.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemolysin_pore.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMs; TIGR01002; hlyII; 1.
SQ SEQUENCE 325 AA; 36958 MW; 19522EA2DA498424 CRC64;

Query Match 27.0%; Score 401; DB 2; Length 325;
Best Local Similarity 28.9%; Pred. No. 7.5e-16;
Matches 81; Conservative 72; Mismatches 107; Indels 20; Gaps 6;

QY 11 TTDIGSNTTGTGDLVTDKENGHMKKVFYSFIDDKHNNKLLVIRTKGTIAGQRYVSE 70
Db 45 TTATSDSKLISQILT-----FNFIDKSYDKDTLILKAAGNIVSGYTKPNP 92
QY 71 EGANKSLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXXMXXTXGXNXXTXDXTX 130
Db 93 KDTISSQFYWGSKYNISINSDSNVVDYAPKN--QNEEFQVQTVGYSG--GDINI 148
QY 131 KXGLXGXN--XSXGTXKXVQDPDKTILESPTD-KYGVKVFNNMWNQNGPYDRDSW 187
Db 149 SNGLSGGGNGSKSFSETINYNKQESYRTSLDKRTNFKKIGWDVEAHKIMNNGWPGYGRDSY 208
QY 188 NPVYGNQLFMKTNGSKMAADNFDLPNKASSLLSGFSDFATVITMDRKSKQQTNDIV 247
Db 209 HSTGNEFMFLGSRNAGQNFLEYHKMPVLVSRGNPEFGLVSRKQNAK-KSKITV 267
QY 248 IYERVDDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWE 287
Db 268 TYQREMDRYTNFWINENIGNYKDNENPATHTSIYEDWE 307

RESULT 4
```



```
FT SIGNAL 1 26 POTENTIAL.
SQ SEQUENCE 327 AA; 36889 MW; 3807DB6421ACDAAA CRC64;
  Query Match 26.7%; Score 395.5; DB 2; Length 327;
  Best Local Similarity 30.0%; Pred. No. 1.6e-15;
  Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKTGDLVTDKENGMMHKVFYFSDIDKNNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
Db 43 TLYKTTATSDNDKLN-ISOILTFNFIKDKSYDKDQTLVLKAAGNINSYGKKPNPKDYNYSQ 101

QY 78 LAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEXMXTYTXGXNKNXTXDYTKXGXGLXG 137
Db 102 FYWGGKYNVSVSSESDAVNVVDYAPKN--QNEEFQVQOTLGSYG--GDINISNGLSGG 157

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVQNGPYDRDSWNPVYGNQ 194
Db 158 LKNGSFSTINYNKQESYRTTIDRKNHKSIGWGVEAHKIMNNGWGPYGRDSYDPTYGNE 217

QY 255 DYQLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 277 RYTNQWNLRLHWIGNNYKNQNTVTFSTVEVDWQ 309

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVQNGPYDRDSWNPVYGNQ 194
Db 158 LKNGSFSTINYNKQESYRTTIDRKNHKSIGWGVEAHKIMNNGWGPYGRDSYDPTYGNE 217

QY 195 LFMKTRNGSMKAADNPLDPNKAASSLLSGFSPDFATVITMDRKASKQQTNDIVYIYVRD 254
Db 218 LFLGGRQSSNAGQNFPLTHQMPLLARGNFPNPEFISVLSHKQNDTK-KSKIKVTYQREMD 276

QY 255 DYQLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 277 RYTNQWNLRLHWIGNNYKNQNTVTFSTVEVDWQ 309

RESULT 7
Q99T54
ID Q99T54 PRELIMINARY; PRT; 327 AA.
AC Q99T54;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Leukotoxin, LukD.
GN LUKD OR SAl637.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J.-Q., Ito T.,
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C.,
RA Hiramatsu K., Sekimizu K.,
RA Yanashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of meticillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003135; BAB42905.1;
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01458; BICOMPOTOXIN.
DR TIGRFAMs; TIGR01002; hlyII; 1.
KW Complete proteome.
SQ SEQUENCE 327 AA; 36903 MW; COA5486421ACDAB9 CRC64;

Query Match 26.6%; Score 394.5; DB 16; Length 327;
Best Local Similarity 30.0%; Pred. No. 1.9e-15;
Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKTGDLVTDKENGMMHKVFYFSDIDKNNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
Db 43 TLYKTTATSDNDKLN-ISOILTFNFIKDKSYDKDQTLVLKAAGNINSYGKKPNPKDYNYSQ 101

QY 78 LAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEXMXTYTXGXNKNXTXDYTKXGXGLXG 137
Db 102 FYWGGKYNVSVSSESDAVNVVDYAPKN--QNEEFQVQOTLGSYG--GDINISNGLSGG 157

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVQNGPYDRDSWNPVYGNQ 194
Db 158 LKNGSFSTINYNKQESYRTTIDRKNHKSIGWGVEAHKIMNNGWGPYGRDSYDPTYGNE 217

QY 195 LFMKTRNGSMKAADNPLDPNKAASSLLSGFSPDFATVITMDRKASKQQTNDIVYIYVRD 254
Db 218 LFLGGRQSSNAGQNFPLTHQMPLLARGNFPNPEFISVLSHKQNDTK-KSKIKVTYQREMD 276

QY 255 DYQLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 277 RYTNQWNLRLHWIGNNYKNQNTVTFSTVEVDWQ 309
```

```
Db 102 FYWGGKYNVSVSSESDAVNVVDYAPKN--QNEEFQVQOTLGSYG--GDINISNGLSGG 157

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVQNGPYDRDSWNPVYGNQ 194
Db 158 LKNGSFSTINYNKQESYRTTIDRKNHKSIGWGVEAHKIMNNGWGPYGRDSYDPTYGNE 217

QY 195 LFMKTRNGSMKAADNPLDPNKAASSLLSGFSPDFATVITMDRKASKQQTNDIVYIYVRD 254
Db 218 LFLGGRQSSNAGQNFPLTHQMPLLARGNFPNPEFISVLSHKQNDTK-KSKIKVTYQREMD 276

QY 255 DYQLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 277 RYTNQWNLRLHWIGNNYKNQNTVTFSTVEVDWQ 309

RESULT 8
Q931N4
ID Q931N4 PRELIMINARY; PRT; 332 AA.
AC Q931N4;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Leukotoxin F-subunit.
GN LUKD OR SAV1819.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Oi Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of meticillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003363; BAB57981.1;
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR Pfam; PF01117; Aerolysin; 1.
DR TIGRFAMs; TIGR01002; hlyII; 1.
KW Complete proteome.
SQ SEQUENCE 332 AA; 37738 MW; 5C896297453A3638 CRC64;

Query Match 26.6%; Score 394.5; DB 16; Length 332;
Best Local Similarity 30.0%; Pred. No. 1.9e-15;
Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKTGDLVTDKENGMMHKVFYFSDIDKNNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
Db 43 TLYKTTATSDNDKLN-ISOILTFNFIKDKSYDKDQTLVLKAAGNINSYGKKPNPKDYNYSQ 101

QY 78 LAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEXMXTYTXGXNKNXTXDYTKXGXGLXG 137
Db 102 FYWGGKYNVSVSSESDAVNVVDYAPKN--QNEEFQVQOTLGSYG--GDINISNGLSGG 157

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVQNGPYDRDSWNPVYGNQ 194
Db 158 LKNGSFSTINYNKQESYRTTIDRKNHKSIGWGVEAHKIMNNGWGPYGRDSYDPTYGNE 217

QY 195 LFMKTRNGSMKAADNPLDPNKAASSLLSGFSPDFATVITMDRKASKQQTNDIVYIYVRD 254
Db 218 LFLGGRQSSNAGQNFPLTHQMPLLARGNFPNPEFISVLSHKQNDTK-KSKIKVTYQREMD 276

QY 255 DYQLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 277 RYTNQWNLRLHWIGNNYKNQNTVTFSTVEVDWQ 309
```

Query Match 26.5%; Score 393; DB 2; Length 325;

```

Best Local Similarity 29.2%; Pred. No. 2.3e-15;
Matches 83; Conservative 72; Mismatches 109; Indels 20; Gaps 6;

Qy 11 TTDTGNTTAVKGLVTVYDKENGHKKVYFSDIDKNNHKLVLVTKGTIAGQYRVYSE 70
Db 47 TTATADSKFKISQILT-----FNIKDKSDKDFLVKAAAGNINSYGERPNP 94

Qy 71 EGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXT--XTXGXNXXNTXDX 128
Db 95 KDYDFSKIYWGAKYNNVSISSQSDSNVVDYAPKQNEEFQVONTLGYTFGGDISNGL 154

Qy 129 TKXGXLGNXNXXGTXKXXVQDPFKTILSPD-KKVGWKVIFNNMVNQNNQYDRDSW 187
Db 155 ---SGGLAG-NTAFSETINYKQSYRTTILSRNTNKNVKGWVEAHKIMNNGWGPYGRDSF 210

Qy 188 NPVYGNQFMKTRNGSKMAADNFDLPNKASSLLSGSPDFATVITMDRKASKQQTNDV 247
Db 211 HPTYNELFLAGRQSSAVAGQNFIAHQHOMPLLSRNFNFPELSVLS-HRQDGAKKSKITV 269

Qy 248 IYERVDDYQLHWTSTNKGNTKDKWTDRRSERYKIDWEKEEM 291
Db 270 TYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 313

RESULT 11
Q931F3 PRELIMINARY; PRT; 325 AA.
AC Q931F3;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Gamma-hemolysin component B.
GN HLCB OR SAV2421.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Goto S., Yabuzaki J.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003365; BAB58583.1;
DR EMBL; AP003365; BAB43092.1;
DR HSSP; P09616; 7AHL.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemolysn_pore.
DR InterPro; IPR003963; Staph_bicn_tcn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPOTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
KW Complete proteome.
SQ SEQUENCE 325 AA; 36658 MW; 082999EB877CA2B5 CRC64;

Query Match 25.9%; Score 385; DB 16; Length 325;
Best Local Similarity 25.9%; Pred. No. 6.7e-15;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

Qy 9 TGTDTI-GSNTTV---KTGDLTVYDKENG-----MHKKVYFSDIDKNNHKLVLVIR 56
Db 21 SGTANAEGKITPVSKKVDKVLTKYTATADSKDKFISQILTTFNFKDKSYDKDTLVLK 80

Qy 57 TKGITAGQYRVYSEGANSGAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXT- 115
Db 81 ATGINSFGVAPNPNDYDFSKLYWGAKYNVSISSQSDSNVVDYAPKQNEEFQVQNTL 140

Qy 116 -XTGXNXXNTXDXTKXGXLGNXNXXGTXKXXVQDPFKTILSPD-KKVGWKVIFNN 173
Db 141 GYTFGGDISNGL---SGGLNG-NTAFSETINYKQSYRTTILSRNTNKNVKGWVEAHK 196

```

```

Qy 174 MVNQWGPYDRDSWNPVYGNQFMKTRNGSKMAADNFDLPNKASSLLSGSPDFATVIT 233
Db 197 IMNNGWGPYGRDSFPHTYGNELFLAGRQSSAVAGQNFIAHQHOMPLLSRNFNFPELSVLS 256

Qy 234 MDRKASKQQTNDVYERVDDYQLHWTSTNKGNTKDKWTDRRSERYKIDWEKEEM 291
Db 257 -HRQDGAKKSKITVTYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 313

RESULT 12
Q99SN8 PRELIMINARY; PRT; 338 AA.
AC Q99SN8;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical protein SAV2004.
GN SAV2004 OR SA1812.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58166.1;
DR EMBL; AP003365; BAB43092.1;
DR HSSP; P09616; 7AHL.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemolysn_pore.
DR InterPro; IPR003963; Staph_bicn_tcn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPOTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 338 AA; 38662 MW; B4B4F335B4B8DB1 CRC64;

Query Match 25.9%; Score 385; DB 16; Length 338;
Best Local Similarity 28.2%; Pred. No. 7.2e-15;
Matches 82; Conservative 81; Mismatches 110; Indels 18; Gaps 8;

Qy 13 DIGSNTTVTKGDLTVYDKENGHKKVYFSDIDKNNHKLVLVIRTKGTIAGQYRVYSE 72
Db 42 NLGDGTRMYTRTATTSQSKNITQSLQFNFLTEPNYDKETVFKAKGTIGSLRIDPNG 101

Qy 73 ANKSGLAWSAFKVLQLPDNEVAQISDYPRNSIDTXEMXTXGXNXXNTXDXTK 131
Db 102 YWNTSLRMPGYSYSIQNVDDNNNTNTVDFAPKQDESREVKT----YGYKTGDFSN 157

Qy 132 XGXLXG---XNKSXGTXKXXVQDPFKTILSPD-KKVGWKVIFNNMVNQNNQYDR 184
Db 158 RGLTGNTKESNTSETISYQPSYRFLLDQSTSHKGVGWKVEAHLNNMGHDTROLTN 217

Qy 185 DSWNPVYGNQFMKTRNGSKMAADNFDLPNKASSLLSGSPDFATVITMDRKASKQQTN 244
Db 218 DSDNRT-KSEIFSLTRNGLWAKNFTPKDKMPVTVSEGFNPEFLAVMSHDK-KDKGSK 275

Qy 245 IDVYERVDDYQLHWTSTN---WKGNTNKTDKWTDRRSERYKIDWEKEEM 291
Db 276 FVHYKRSMDDEFKIDWNRHGFNGWYSGENHVDKKEELSLALYEDWKTHDV 326

```

```
RESULT 13
Q54327
ID Q54327 PRELIMINARY; PRT; 326 AA.
AC Q54327;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Synergohymenotrophic toxin.
GN LukF.1.
OS Staphylococcus intermedius.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B62;
RX MEDLINE=96105366; PubMed=7498527;
RA Prevost G., Bouakham T., Piemont Y., Monteil H.;
RT "Characterisation of a synergohymenotrophic toxin produced by
Staphylococcus intermedius.";
RL FEBS Lett. 376:135-140(1995).
DR EMBL: X79188; CAA55783.1; -.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR Pfam; PF01117; Aerolysin.1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
DR CHAIN 27 326 SYNERGOHYMENOTROPIC TOXIN.
SQ SEQUENCE 326 AA; 36597 MW; F8A15CE8345A22B4 CRC64;

Query Match 25.9%; Score 384.5; DB 2; Length 326;
Best Local Similarity 30.1%; Pred. No. 7.3e-15;
Matches 83; Conservative 72; Mismatches 108; Indels 13; Gaps 9;

QY 18 TTVTGTDLVTVYDREN-GMHKKVYFSFIDDKHNHKKLLIVRTKGTIAGQYRVYSEEGANKS 76
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
D 43 TLYKT--TATADSDNLNISQLLFNFKDKSYDKDILVLKAAGNINSYKSPNPDIYS 100
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 77 GLAWPAFKVOLQPDNEVAQISDYPRNSIDTXEMXMYTXGXNXXNXTXDTXXGXGX 136
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 101 SFYWGAKYVNSISAESKGAVNVVDYAPKN--QNEEFQVQNTLCYSFG--GDISISKGLSG 156
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 137 GXN--XSXGXTXXKXVOPDEKFTILESPTDK-VGKWKVIFNNMVNQNPYDRDSWNPVYGN 193
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D 157 GLNGSEFSETINIKQESYRTTIDKHTDNKTIGWGVEAHKIMNAGWPGYGRDSFHDLYGN 216
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 194 GLPMKTRNGSKAADNFDLPNKASSLSG-FSPDFATVITMDRKASKQQTNDIVYERV 252
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D 217 ELFLGGRQSKLNAGQNL-PTROMPLARGNFPFLSVLSHKNPKGAK-TSKIKVYQRE 274
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 253 RDYQLHWTSTNKGNTKDKWTDSSERYKIDWEK 288
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D 275 MDEYTYWNGFHWGNTGNYKNQNNATFTSFYEIDWDQ 310
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 14
Q53704
ID Q53704 PRELIMINARY; PRT; 322 AA.
AC Q53704;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE LukF-pv like component.
GN LukF-PV.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=P83;
RA Kaneko J., Choorit W., Kamio Y.;
RT "LukF-pv like component of leukocidin and gamma-hemolysin in S.aureus
p8 3.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: D83951; BAAL2148.1; -.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin.1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
DR SEQUENCE 322 AA; 36519 MW; 9424E58715B108E6 CRC64;

Query Match 25.2%; Score 374; DB 2; Length 322;
Best Local Similarity 27.5%; Pred. No. 3e-14;
Matches 84; Conservative 75; Mismatches 119; Indels 28; Gaps 6;

QY 6 NIKTGTDTIGSNTVTGDLVTVYDKENGHHKV-----FYSFIDD 45
| : | | | : | | : | | : | : | : | : | : | : | : | : | : | : |
D 10 SVATISITLIMLSNTVDAQAQHTPVSEKKVDDKITLYKTATSDSKLKISQILTFNFIK 69
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 46 KHNHKKLLIVRTKGTIAGQYRVYSEEGANKSGLAWPAFKVOLQPDNEVAQISDYPRN 105
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D 70 KSYDKDTLILKAAAGNIYSYGTQPTSDSSINSQFYWGAKNVYVSSSEKSDSVNVDYAPKN 129
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 106 SIDDTEXMXT--XTGXKNXNXTXDTXKXGLXGXNKSXGTXKXVQDPDKTILESPD- 162
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D 130 QNEEFQVQNTLCYSYGGDINIINGLT---GGLNG-SKSFSETINIKQESYRTTIDRKTNH 185
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 163 KYGVKVIFFNNMVNQNPYDRDSWNPVYGNQNLFMKTRNGSKAADNFDLPNKASSLS 222
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D 186 KSIGWGVEAHKIMNNGWPGYGRDSSDSLYGNELFLGGRQSSNANQNFLPTHQMPILARG 245
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 223 GFSPDFATVITMDRKASKQQTNDIVYERVDDYQLHWTSTNKGNTKDKWTDSSERY 282
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D 246 NFNPESVLSHKKQDKV-KSIKVTYQREMDRYENFWNNLHWIGYNIKNKQKRATHTSY 304
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 283 KIDWEK 288
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
D 305 EIDWEK 310
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 15
Q9MBN2
ID Q9MBN2 PRELIMINARY; PRT; 322 AA.
AC Q9MBN2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE LukF-PV(p83).
OS Staphylococcus aureus prophage phiPV83.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=129009;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P83;
RA Zou D., Kaneko J., Narita S., Kamio Y.;
RT "Complete nucleotide sequence and molecular characterization of
prophage PV83pro carrying lukM-lukF-pv(p83) gene cluster in
Staphylococcus aureus strain p83.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=P83;
RX MEDLINE=98067870; PubMed=9404084;
RA Kaneko J., Muramoto K., Kamio Y.;
RT "Gene of LukF-pv-like component of Panton-Valentine leukocidin in
Staphylococcus aureus p83 is linked with lukM.";
RL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).
DR EMBL: AB044554; BAA97867.1; -.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
```

DR Pfam: PF01117; Aerolysin; 1.
DR TIGRFAMs: TIGR01002; hlyII; 1.
SQ SEQUENCE 322 AA; 36519 MW; 9424E58715B108E6 CRC64;
Query Match 25.28; Score 374; DB 9; Length 322;
Best Local Similarity 27.5%; Pred. No. 3e-14;
Matches 84; Conservative 75; Mismatches 119; Indels 28; Gaps 6;
Qy 6 NIKGTDTDIGSNTVVKTGDLVYDKENGMHKKV-----FYSFIDD 45
Db 10 SVATSIILMLNSNTVDAQAHTVPSEKKVDDKITLYKTATSDSKLKISQILTFNFIKD 69
Qy 46 KNHNKKLIVIRTKGTAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDIYPRN 105
Db 70 KSYDKDTLILKAAGNIYSGYTQPTSDSSINSQFYWGAKYNVFSSEKDSVNIYDAPKN 129
Qy 106 SIDTXEMXT--XTXGXNXXNXTDXTKXGXKXNXXSGXTXKXVOPDFKTTILESPTD- 162
Db 130 QNEEFVQQTTLGYSYGGDINIINGLT---GGLNG-SKSFSETINVKQESYRTTIDRKTNH 185
Qy 163 KKVGVKVIENMVNQNMGPDYDRDSNPNVYGNLFEMKTEGSMKAADNFDPNKASSLLSS 222
Db 186 KSIQGVFAHKIMNMGWGPYGRDSDSLYGNELFGLGGQSSNANQNPFLPHQMPILARG 245
Qy 223 GFSPDFATVITMDRKASKQQTNIIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDRSERY 282
Db 246 NFNPEFISVLSHKQKDYK-KSKIKVTYQRENDRYENFNWNLHWIGYNIKNQKRATHTSIY 304
Qy 283 KIDWEK 288
Db 305 EIDWEK 310

Search completed: March 4, 2003, 10:44:08
Job time : 30.5 secs